PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 16/00, C12N 15/13 — A2

(11) International Publication Number: WO 95/15982

(43) International Publication Date: 15 June 1995 (15.06.95)

(21) International Application Number: PCT/US94/14106

(22) International Filing Date: 8 December 1994 (08.12.94)

(30) Priority Data:

08/164,022	8 December 1993 (08.12.93)	US
08/350,400	6 December 1994 (06.12.94)	US

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published

Without international search report and to be republished upon receipt of that report.

(71) Applicant: GENZYME CORPORATION [US/US]; One Kendall Square, Cambridge, MA 02139 (US).

(72) Inventors: BARSOMIAN, Gary; 68 Jewett Street, Georgetown, MA 01830 (US). COPELAND, Diane, P.; 7 Whittier Road, Billerica, MA 01821 (US). HILLHOUSE, Dana; Top floor, 33 Whitney Avenue, Beverly, MA 01915 (US). JOHNSON, Tracy; 136 Concord Avenue, Lexington, MA 02173 (US).

(74) Agents: VINCENT, Matthew, P. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).

(54) Title: PROCESS FOR GENERATING SPECIFIC ANTIBODIES

(57) Abstract

The present invention relates to a method for generating an antibody which is specific for an immunorecessive epitope, and nucleic acid encoding the antibody. The subject method generally comprises the steps of generating a variegated display library of antibody variable regions, and selecting from the library those antibody variable regions which have a desired binding specificity for the immunorecessive epitope. The antibody variable regions used to generate the display library are cloned from an immunotolerance-derived antibody repertoire.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

Process for Generating Specific Antibodies

Background of the Invention

In an antibody-producing animal, such as a mammal, antibodies are synthesized and secreted into bodily fluids by plasma cells, a type of terminally differentiated B-lymphocyte. Exposure of the animal to a foreign molecule (i.e. via immunization) generally produces multiple plasma cell clones resulting in a heterogeneous mixture of antibodies (polyclonal antibodies) in the blood and other fluids. The blood of an immunized animal can be collected, clotted, and the clot removed to leave a sera containing the antibodies produced in response to immunization. This remaining liquid or serum, which contains the polyclonal antibodies, is referred to as antiserum. However, such antiserum contains many different types of antibodies that are specific for many different antigens. Even in hyperimmunized animals, seldom are more than one tenth of the circulating antibodies specific for the particular immunogen used to immunized the animal. The use of these mixed populations of antibodies, though useful in many situations, can create a variety of different problems in immunochemical techniques. For example, such antiserum will generally be inadequate for use in distinguishing between the immunogen and closely related molecules which share many common determinants with the immunogen.

Owing to their high specificity for a given antigen, the advent of monoclonal antibodies (Kohler and Milstein (1975) *Nature* 256:495) represented a significant technical break-through with important consequences both scientifically and commercially. Monoclonal antibodies (MAbs) are traditionally made by isolating a single antibody secreting cell (e.g. a lymphocyte) from an immunized animal, fusing the lymphocyte with a myeloma (or other immortal) cell to form a hybrid cell (called a "hybridoma"), and then culturing the selected hybridoma cell *in vivo* or *in vitro* to yield antibodies which are identical in structure and specificity. Because the antibody-secreting cell line is immortal, the characteristics of the antibody are reproducible from batch to batch. The usefulness of monoclonal antibodies stems from three characteristics - their specificity of binding, their homogeneity, and their ability to be produced in virtually unlimited quantities.

While production of monoclonal antibodies has resulted in production of antibodies of greater specificity to a particular antigen then polyclonal methods, there are nevertheless a number of limitations associated with these techniques and antibodies produced thereby. For instance, a key aspect in the isolation of monoclonal antibodies relates to how many antibody producing hybridoma cells with different specificities can be practically established and sampled in response to immunization with a particular antigen, compared to how many theoretically need to be sampled in order to obtain an antibody having specific characteristics.

For example, the number of different antibody specificities expressed at any one time by lymphocytes of the murine immune system is thought to be approximately 10^7 and represents only a small proportion of the potential repertoire of specificities.

Immunization regimens can provide enrichment of B-cells producing the desired antibodies. However, even employing those techniques, typical protocols for isolating antibody producing B-cells permit sampling of generally less than 500 antibody producing hybridoma cells per immunized animal. Thus, traditional techniques for the production of monoclonal antibodies statistically favor generation of monoclonal antibodies to immunodominant molecules, making isolation of antibodies specific for a rare or less immunodominant epitope difficult. This problem can be further exacerbated by the fact that in many instances pure antigen is not available as an immunogen, particularly in the case of cell surface antigens. Immunization with intact cells frequently results in production of antibodies against irrelevant epitopes, especially for xenotypic immunization. To enhance the production of monoclonal antibodies to rare, "immunorecessive" antigens, immunotolerance techniques have been employed. Neonatal tolerization and chemical immunosuppression are most commonly used to reduce clonal expansion of B cells in response to "background" antigen signals, thereby enriching for a population of B cells responsive to the epitopes of interest. However, the practical application of a subtractive immunization technique can be very difficult, as the efficiency of immunosuppression is often not acceptable, or as in the case of cyclophosphamide immunosuppression, generally results in only a few antibody-producing hybridoma cells per immunotolerized animal (e.g. less than 100), making it unlikely that monoclonal antibodies can be isolated which are specific to the immunorecessive epitopes.

Summary of the Invention

The present invention provides a method for generating an antibody which is specific for an immunorecessive epitope, and nucleic acid encoding the antibody. The subject method generally comprises the steps of generating a variegated display library of antibody variable regions, and selecting from the library those antibody variable regions which have a desired binding specificity for the immunorecessive epitope. The antibody variable regions used to generate the display library are cloned from an immunotolerance-derived antibody repertoire.

As described herein, the antibody variable regions of the display library are presented by a replicable genetic display package in an immunoreactive context which permits the antibody to bind to an antigen that is contacted with the display package. Thus, affinity selection techniques can be utilized to enrich the population of display packages for those having antibody variable regions which have a desired binding specificity for the

immunorecessive epitope. In exemplary embodiments, the display library can be a phage display library. Alternatively, the display library can be generated on a bacterial cell-surface or a spore.

The subject method can be used to isolate antibodies which are specific for such immunorecessive epitopes as, for example, cell-type specific markers, including fetal cell markers such as fetal nucleated red blood markers, cancer cell markers such as colon cancer markers or metastatic tumor cell markers, stem cell markers such as markers for precursor nerve cells or hematopoietic stem cells.

Likewise, the subject method can be used to generate antibodies which can discriminate by binding between a variant form of a protein and other related forms of the protein. The variant protein can differ by one or more amino acid residues from other related proteins in order to give rise to the immunorecessive epitope, as well as vary antigenically from the related protein by virtue of glycosylation or other post-translational modification. The variation can arise naturally, as between different isoforms of a protein family, illustrated by the apolipoprotein E family, or can be generated by genetic aberration, as illustrated by the neoplastic transforming mutations of oncogenic proteins or tumor suppressor proteins such as p53.

In an illustrative embodiment of the subject method, a specific antibody to an immunorecessive epitope can be generated by affinity purification of a antibody phage display library derived from an immunotolerance-derived antibody repertoire. For example, suitable host cells are transformed with a library of replicable phage vectors encoding a library of phage particles displaying a fusion antibody/coat protein, where the fusion protein includes a phage coat protein portion and an antibody variable region portion. The antibody variable region is obtained from the immunotolerance-derived antibody repertoire. The transformed cells are cultured, the phage particles are formed, and the antibody fusion proteins are expressed. Any of resulting phage particles which have an antibody variable region portion which specifically binds to a an immunorecessive epitope can be separated from those which do not specifically bind the immunorecessive epitope.

The present invention further pertains to novel immunorecessive antibody libraries produced by the subject method. From the subject method, for example, an antibody display library can be isolated which is enriched for antibodies that specifically bind an immunorecessive epitope of interest. The display library comprises a population of display packages expressing a variegated V-gene library which has been cloned from an immunotolerance-derived antibody repertoire, and which has been further enriched after expression by the display package via affinity separation with the immunorecessive epitope.

It is also contemplated by the present invention that individual antibodies, and genes encoding these antibodies, can be isolated from the antibody libraries of the subject method. For instance, after affinity enrichment of the antibody display library for antibodies which specifically bind the immunorecessive epitope, individual display packages can be obtained, and the antibody gene contained therein subcloned into other appropriate expression vectors suitable for production of the antibody for the desired use.

Description of the Drawings

Figures 1A and 1B show variable region PCR primers for amplifying the variable regions of both heavy and light chains from murine antibody genes.

Figure 2 shows a schematic representation of an Fab' expression cassette.

Figure 3 is a semi-log graph depicting the binding of phage antibody pools (phab) enriched on the HEL cell line (number indicates the round of enrichment). The graph provides additional comparison of the enriched phab pools with the binding of other immunoglobulins (T3, Anti-M and Wilma) to the HEL cells.

Figure 4 illustrates the percentage of cells (either HEL cells or mature white cells) stained by individual phab isolates generated by the subject method.

Figure 5A shows the results of sequential rounds of pre-adsorption and enrichment on fetal liver cells for phab binding. The increase in the percentage of phage antibodies binding to fetal liver cells is indicative enrichment for fetal cell binding phage antibodies. The phage antibody library was derived using a V-gene library from an immunotolerized host animal. In contrast, Figure 5B compares the results of the immunotolerized experiment in Figure 5A with the results of sequential rounds of panning using phage antibody libraries derived immunized, but not tolerized, host animals.

Figure 6 show variable region PCR primers for amplifying the variable regions of both heavy and light chains from human antibody genes.

Figure 7 details the sequences for CDR3 regions of both heavy and light chains for individual phab isolates enriched on fetal cells.

Figures 8A and 8B illustrate the general features of the FB3-2 and H3-3 antibodies, respectively, including the framework regions (double underline; FRs), complementarity determining regions (CDRs), and constant regions (italics; IgG1 CH1 or kappa constant). The amino acid residues which differ between the FB3-2 and F4-7 antibodies are indicated under the FB3-2 sequence in Figure 8A.

Detailed Description of the Invention

The present invention makes available a powerful directed approach for isolating specific antibodies which are extremely difficult or impossible to obtain by current methodologies, and thereby overcomes the deficiencies discussed above. One aspect of the present invention is the synthesis of a method that combines immunotolerization and variegated display libraries to yield a dramatic and surprising synergism in the efficient isolation of antibodies having a desired binding affinity for an immunorecessive target epitope. Utilizing immunotolerance techniques such as subtractive immunization, a subset of lymphocytes producing antibodies against an immunorecessive target epitope are enriched in an immunized animal. Subsequent isolation of antibody-producing cells from the immunized animal and PCR amplification of at least the variable regions of antibodies expressed by the isolated cells allows the generation of a variegated library of antibody variable region genes (V-genes). From this V-gene library, the subject method selects genes encoding antibodies specific for the target epitope by (i) displaying the antibodies encoded by each variable region gene on the outer surface of a replicable genetic display package to create an antibody display library, and (ii) using affinity selection techniques to enrich the population of display packages for those containing V-genes encoding antibodies which have a desired binding specificity for the target epitope.

In general, most antibodies isolated by recombinant antibody display technologies known in the art are obtained using substantially pure preparations of an antigen of interest, and provide only a few isolates having association constants (K_as) even approaching 10^9M^{-1}. No phage display method has resulted in isolation of antibodies panning with live cells (i.e., unpurified antigen) which are of the equivalence in either specificity or affinity to antibodies attainable by conventional hybridoma techniques. In contrast, as demonstrated in the Examples provided below, the subject method can be used to generate antibodies which out perform both the combinatorial and hybridoma-derived antibodies of the prior art, particularly with respect to binding affinity and degrees of specificity.

For example, antibodies isolated by the subject method can have binding affinities greater than 10^8M^{-1}, e.g., in the range of 10^9M^{-1} to $10^{12}M^{-1}$. Moreover, the specificity of these antibodies can be several fold, if not orders of magnitude, better than combinatorial and hybridoma generated antibodies, particularly with respect to antibodies for cell surface epitopes. For instance, the subject method can provide antibodies which have no substantial background binding to other related cells, e.g., specificities greater than 10 fold binding to the target cells over background binding to the related cells. As demonstrated below, antibodies can be generated which do not substantially cross-react with other epitopes, preferably having

specificities greater than 20 fold over background, more preferably 50, 75 or 100 fold over background, and even more preferably more than 125 fold over background.

For the purpose of the present invention, the term "antibody" in its various grammatical forms is art-recognized and includes immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen. Structurally, the simplest naturally occurring antibody (IgG) comprises four polypeptide chains, two heavy (H) chains and two light (L) chains inter-connected by disulfide bonds. The light chains exist in two distinct forms called kappa (κ) and lambda (λ). Each chain has a constant region (C) and a variable region (V). Each chain is organized into a series of domains. The light chains have two domains, corresponding to the C region and the other to the V region. The heavy chains have four domains, one corresponding to the V region and three domains (1,2 and 3) in the C region. The naturally occurring antibody has two arms (each arm being an Fab region), each of which comprises a V_L and a V_H region associated with each other. It is this pair of V regions (V_L and V_H) that differ from one antibody to another (owing to amino acid sequence variations). The variable domains for each of the heavy and light chains have the same general structure, including four framework regions (FRs), whose sequences are relatively conserved, connected by three hypervariable or complementarity determining regions (CDRs). The variable region of each chain can typically be represented by the general formula FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4. The CDRs for a particular variable region are held in close proximity to one and other by the framework regions, and with the CDRs from the other chain and which together are responsible for recognizing the antigen and providing an antigen binding site (ABS).

Moreover, it has been shown that the function of binding antigens can be performed by fragments of a naturally-occurring antibody, and as set out above, these antigen-binding fragments are also intended to be designated by the term "antibody". Examples of binding fragments encompassed within the term antibody include (i) the Fab fragment consisting of the V_L, V_H, C_L and C_{H1} domains; (ii) the Fd fragment consisting of the V_H and C_{H1} domains; (iii) the Fv fragment consisting of the V_L and V_H domains of a single arm of an antibody, (iv) the dAb fragment (Ward et al., (1989) *Nature* 341:544-546) which consists of a V_H domain; (v) isolated CDR regions; and (vi) $F(ab')_2$ fragments, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region. Furthermore, although the two domains of the Fv fragment are coded for by separate genes, it has proved possible to make a synthetic linker that enables them to be made as a single protein chain (known as single chain Fv (scFv); Bird et al. (1988) *Science* 242:423-426; and Huston et al. (1988) *PNAS* 85:5879-5883) by recombinant methods. Such single chain antibodies are also encompassed within the present meaning of the term "antibody".

The language "antibody variable region" is likewise recognized in the art, and includes those portions of an antibody which can assemble to form an antigen binding site. For instance, an antibody variable region can comprise each of the framework regions (FR1-FR4) and complementary determining regions (CDR1-CDR3) for one or both chains of an IgG molecule.

The language "a desired binding specificity for an immunorecessive epitope", as well as the more general language "antibody specificity", refers to the ability of individual antibodies to specifically immunoreact with distinct antigens. The desired binding specificity will typically be determined from the reference point of the ability of the antibody to differentially bind, and therefore distinguish between, two different antigens -particularly where the two antigens have unique epitopes which are present along with many common epitopes. For instance, a desired binding affinity for an immunorecessive epitope can refer to the ability of an antibody to distinguish between related cells, such as between adult and fetal cells, or between normal and transformed cells. In other embodiments, the desired binding affinity can refer to the ability of the antibody to differentially bind a mutant form of a protein versus the wild-type protein, or alternatively, to discriminate in binding between different isoforms of a protein. An antibody which binds specifically to an immunorecessive epitope is referred to as a "specific antibody". The term "relative specificity" refers to the ratio of specific immunoreactivity to background immunoreactivity (e.g., binding to non-target antigens). For instance, relative specificity for fetal cells can be expressed as the ratio of the percent binding to fetal cells to the percent binding to maternal cells. Antibodies which have no substantial background binding to a non-target antigen, such as a maternal cell, have large relative specificities (e.g., in excess of 10 fold over background binding).

The phrases "individually selective manner" and "individually selective binding", with respect to immunoreactivity of an antibody with a particular cell, refers to the binding of an antibody to a certain cell phenotype which binding, in addition to being phenotypically dependent, is also dependent on the particular individual from which the cell is isolated, e.g. the source of the cell. Individually selective binding does not refer to inter-species specificity of binding, rather relates to intraspecies specificity.

Antibody binding to antigen, though entirely non-covalent, can nevertheless be exquisitely specific for one antigen versus another, and often very strong. Antibodies can specifically bind different structural components of most complex protein, nucleic acid, and polysaccharide antigens. In general, macromolecules are much bigger than the antigen binding site of an antibody. Therefore, an antibody binds to only a particular portion of the macromolecule, referred to herein as the "determinant" or "epitope". The total number of antibodies produced by a population of antibody-producing cells in a particular animal is referred to a the "antibody repertoire". The extraordinary diversity of the antibody repertoire

is a result of variability in the structures of the antigen binding sites amongst the individual antibodies which make up the repertoire.

The process of "immunization" refers to the exposure of an animal (that is capable of producing antibodies) to a foreign antigen so as to induce active immunity, which includes the production of antibodies to the foreign antigen. Molecules that generate an immune response are called immunogens.

The language "immunorecessive epitope", which is also substituted from time to time with the terms "rare epitope" or "target epitope", is intended to refer to epitopes that, in the context that it ordinarily occurs or can be isolated as an immunogen, are typically not efficient for use in generating an antibody response by immunization, at least so far as polyclonal and monoclonal antibody production is concerned. Such immunorecessive epitopes will generally be less abundant and/or less antigenic than other epitopes commonly associated with them in the immunogen. Even under circumstances wherein the immunorecessive epitope can elicit a strong antibody response, this response can be, for example, statistically masked by the overall number of antibodies produced as a consequence of the antigenic challenge due to other epitopes associated with the immunorecessive epitope in the immunogen (referred to herein as "immunodominant epitopes" or "background epitopes"). Immunorecessive epitopes may be associated with, for example, cell surface antigens that are unique to a particular cell phenotype. In many instances, this cell surface antigen is not in and of itself available as an immunogen because no purified form of the antigen has been obtained. This can be especially true in the instance of integral membrane proteins that lose their conformation during purification. Thus, an immunogen containing the immunorecessive epitope will also include many background epitopes which can act to decrease the overall percentage of B-lymphocytes activated by the immunorecessive epitope in the total B-lymphocyte population. In an exemplary embodiment, the immunogen can comprise the whole cell on which the immunorecessive epitope is expressed. For example, the immunorecessive epitope can be a cell-type specific marker, such as a cancer cell marker, a fetal cell marker, or a stem cell marker. Likewise, an immunorecessive epitope can comprise an epitope unique to a variant form of a protein, such as a variant which differs by only one or two amino acid residues from a related protein. For instance, the immunorecessive epitope can be a determinant of a mutant p53 which does not arise on the wild-type p53, or an epitope which unique to a particular isoform of human apolipoprotein E, such as ApoE4.

"Tolerization" refers to the process of diminishing an animal's immunological responsiveness to a potentially antigenic substance present in that animal, and the antigenic substance to which tolerance is created is refered to as a "toleragen". Tolerance results from the interaction of toleragen with antigen receptors on lymphocytes under conditions in which

the lymphocytes, instead of becoming activated, are killed or rendered unresponsive. Tolerance to particular antigens, or more exactly, to particular epitopes of an antigen, can be induced by a number of means, including neonatal tolerization or chemically-induced tolerization, and can be the result of induced clonal deletion or clonal anergy. The route of administration of an antigen can also effect the ability of the antigen to act as either an immunogen or as a toleragen.

The language "immunotolerizing means" relates to a process whereby the antibody response to an immunorecessive epitope is unmasked by the deletion of an antibody response to the background epitopes. For instance, as a first step in the immunotolerizing means, an animal is exposed to a toleragen comprising the immunodominant epitopes. The toleragen, however, lacks the immunorecessive epitopes. After tolerance to these background epitopes has been induced, an immunogen which includes the immunorecessive epitopes, is administered to the animal. Due to the deletion of the antibody response to the background epitopes, the percentage of B-cells activated in response to the rare epitopes are increased relative to the total B-cell population of the animal. That is, the immunotolerizing means can be used to "enrich" for cells producing antibodies specific for an immunorecessive epitope. Thus, as used herein, the term "background epitopes" is further defined as those epitopes that are common between the immunogen and the toleragen, while the term "immunorecessive epitopes" is further understood to refer to epitopes unique to the immunogen (relative to the toleragen). The immunogen and the toleragen will typically be closely related, as for example, in the instance of phenotypically related cells, or mutant or different isoforms of a protein.

The language "immunotolerance-derived antibody repertoire" refers to the population of antibody-producing cells, and their antibodies, generated by an immunotolerization which is intended to enrich for antibodies for an immunorecessive epitope.

The language "variegated V-gene library" refers to a mixture of recombinant nucleic acid molecules encoding at least the antibody variable regions of one or both of the heavy and light chains of the immunotolerance-derived antibody repertoire. A population of display packages into which the variegated V-gene library has been cloned and expressed on the surface thereof is likewise said to be a "variegated antibody display library" or "antibody display library".

The language "replicable genetic display package" or "display package" describes a biological particle which has genetic information providing the particle with the ability to replicate. The package can display a fusion protein including an antibody derived from the variegated V-gene library. The antibody portion of the fusion protein is presented by the display package in an immunoreactive context which permits the antibody to bind to an

antigen that is contacted with the display package. The display package will generally be derived from a system that allows the sampling of very large variegated V-gene libraries, as well as easy isolation of the recombinant V-genes from purified display packages. The display package can be, for example, derived from vegetative bacterial cells, bacterial spores, and bacterial viruses (especially DNA viruses). A variegated mixture of display packages encoding at least a portion of the V-gene library is also referred to as an "antibody display library".

The language "differential binding means", as well as "affinity selection" and "affinity enrichment", refer to the separation of members of the antibody display library based on the differing abilities of antibodies on the surface of each of the display packages of the library to bind to the target epitope. The differential binding of an immunorecessive epitope by antibodies of the display can be used in the affinity separation of antibodies which specifically bind the immunorecessive epitope from antibodies which do not. For example, the same molecule or cell that was used as an immunogen in the immunotolerizing step can also be used in an affinity enrichment step to retrieve display packages expressing antibodies which specifically bind it. Typically, the affinity selection protocol will also include a pre-enrichment step wherein display packages capable of specifically binding the background epitopes are removed. Examples of affinity selection means include affinity chromatography, immunoprecipitation, fluorescence activated cell sorting, agglutination, and plaque lifts. As described below, the affinity chromatography includes bio-panning techniques using either purified, immobilized antigen as well as whole cells.

In an exemplary embodiment of the present invention, the display package is a phage particle which comprises an antibody fusion coat protein that includes the amino acid sequence of an antibody variable region from the variegated V-gene library. Thus, a library of replicable phage vectors, especially phagemids (as defined herein), encoding a library of antibody fusion coat proteins is generated and used to transform suitable host cells. Phage particles formed from the chimeric protein can be separated by affinity selection based on the ability of the antibody associated with a particular phage particle to specifically bind a target epitope. In a preferred embodiment, each individual phage particle of the library includes a copy of the corresponding phagemid encoding the antibody fusion coat protein displayed on the surface of that package. Purification of phage particles based on the ability of an antibody displayed on an individual particle to bind a particular epitope therefore also provides for isolation of the V-gene encoding that antibody. Exemplary phage for generating the present variegated antibody libraries include M13, f1, fd, If1, Ike, Xf, Pf1, Pf3, λ, T4, T7, P2, P4, ϕX-174, MS2 and f2.

The language "fusion protein" and "chimeric protein" are art-recognized terms which are used interchangeably herein, and include contiguous polypeptides comprising a first

polypeptide covalently linked via an amide bond to one or more amino acid sequences which define polypeptide domains that are foreign to and not substantially homologous with any domain of the first polypeptide. One polypeptide from which the fusion protein is constructed comprises a recombinant antibody derived from the cloned V-gene library. A second polypeptide portion of the fusion protein is typically derived from an outer surface protein or display anchor protein which directs the "display package" (as hereafter defined) to associate the antibody with its outer surface. As described below, where the display package is a phage, this anchor protein can be derived from a surface protein native to the genetic package, such as a viral coat protein. Where the fusion protein comprises a viral coat protein and an antibody it will be referred to as an "antibody fusion coat protein". The fusion protein may further comprise a signal sequence, which is a short length of amino acid sequence at the amino terminal end of the fusion protein, that directs at least a portion of the fusion protein to be secreted from the cytosol of a cell and localized on the extracellular side of the cell membrane.

Gene constructs encoding fusion proteins are likewise referred to a "chimeric genes" or "fusion genes".

The term "chimeric antibody" is used to describe a protein including at least the antigen binding portion of an immunoglobulin molecule attached by peptide linkage to at least a part of another protein. A chimeric antibody can be, for example, an interspecies chimera, having a variable region derived from a first species (e.g. a rodent) and a constant region derived from a second species (e.g. a human), or alternatively, having CDRs derived from a first species and FRs and a constant region from a second species.

The term "vector" refers to a DNA molecule, capable of replication in a host cell, into which a gene can be inserted to construct a recombinant DNA molecule.

The terms "phage vector" and "phagemid" are art-recognized and generally refer to a vector derived by modification of a phage genome, containing an origin of replication for a bacteriophage, and preferably, though optional, and origin for a bacterial plasmid. The use of phage vectors rather than the phage genome itself provides greater flexibility to vary the ratio of chimeric antibody/coat protein to wild-type coat protein, as well as supplement the phage genes with additional genes encoding other variable regions, such as may be useful in the two chain antibody constructs described below.

The language "helper phage" describes a phage which is used to infect cells containing a defective phage genome or phage vector and which functions to complement the defect. The defect can be one which results from removal or inactivation of phage genomic sequence required for production of phage particles. Examples of helper phage are M13K07, and M13K07 gene III no. 3.

The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the macromolecule. For example, an isolated nucleic acid encoding one of the subject antibodies preferably includes no more than 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the antibodies gene in genomic DNA, more preferably no more than 5kb of such naturally occurring flanking sequence. The term isolated as used herein also refers to a nucleic acid or peptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state.

In one aspect, the subject invention sets forth a method for rapid and efficient isolation of cell-type specific antibodies. For example, antibodies that specifically bind epitopes unique to fetal cells or, alternatively, epitopes unique to cancer cells, can be generated by the subject method. Likewise, the subject method can be employed to generate antibodies to variant forms of a protein, and which can be used, for example, to detect a mutation of a protein or to differentiate amongst various isoforms of a protein. Thus, the present invention can provide antibodies useful for purification, diagnostic, and therapeutic applications.

In another aspect, the invention concerns novels immunorecessive antibody libraries produced by the subject method, as well as individual antibodies isolated therefrom. From the subject method, for example, an antibody display library can be isolated which is enriched for antibodies that specifically bind an immunorecessive epitope of interest. The display library comprises a population of display packages expressing a variegated V-gene library which has been cloned from an immunotolerance-derived antibody repertoire, and which has been further enriched after expression by the display package by affinity separation with the immunorecessive epitope. Thus, antibody display libraries can be generated which are enriched for specific antibodies to cell surface markers, such as fetal cell of tumor cell markers, as well as variant forms of proteins.

As the immunorecessive epitope is dependent on the difference between the immunogen and toleragen used to generate the immunotolerance-derived antibody repertoire, the specificity of the antibodies enriched for in the subject library can be defined in terms of the particular immunogen/toleragen sets used. For example, where the specific antibody is desired for distinguishing between various cells of common or similar origin or phenotype, the cell to which a specific antibody is desired is used as the immunogen, while a related cell(s) from which it is to be distinguished is employed as the toleragen. Cell-type specific markers for the cell of interest are represented in the immunorecessive epitopes. To illustrate,

wherein the cell-type specific marker is a marker for fetal nucleated red blood cells, the toleragen can include maternal erythroid cells and the immunogen can be fetal erythroid cells. Likewise, where the marker is for colon cancer, the toleragen can comprise normal colon cells and the immunogen can be selected from a colon carcinoma cell line. Other exemplary immunogen/toleragen sets useful for generating the subject antibody libraries, as well as individual antibodies therefrom, are provided in the following description and others will be apparent to those skilled in the art.

Similarly, by choice of the immunogen/toleragen sets, the subject libraries can be generated so as to be enriched for specific antibodies able to distinguish by binding between a variant form of a protein and other related forms of the protein. The variant protein can differ by one or more amino acid residues from other related proteins in order to give rise to the immunorecessive epitope, as well as vary antigenically from the toleragen by virtue of glycosylation or other post-translational modification. The variation can arise naturally, as between different isoforms of a protein family, illustrated by the apolipoprotein E family, or can be generated by genetic aberration, as illustrated by the neoplastic transforming mutations of oncogenic proteins or tumor suppressor proteins such as p53.

It is also contemplated by the present invention that individual antibodies, and genes encoding these antibodies, can be isolated from the antibody libraries of the subject method. For instance, after affinity enrichment of the antibody display library for antibodies which specifically bind the immunorecessive epitope, individual display packages can be obtained, and the antibody gene contained therein subcloned into other appropriate expression vectors suitable for production of the antibody for the desired use.

The major aspects of the subject invention will be generally described below and preferred embodiments will be more specifically described in the attached examples.

I. Immunotolerization

Immunotolerization can be employed in the present invention to generate an antibody repertoire, for use in subsequent V-gene cloning steps, in which the antibody response to an immunorecessive epitope(s) has been unmasked. Immunotolerization can be carried out in either *in vivo* or *in vitro* immunization systems. For instance, immunotolerization can be employed in the present invention to enrich the pool of activated B-lymphocytes in an immunized animal for cells producing antibodies directed to immunorecessive epitopes of interest. In a typical immunotolerization procedure of the subject method, an immunogen is introduced to the immune system of an animal some time after exposure to a toleragen. The effect of the toleragen is to reduce or abrogate altogether any immunological response upon

re-exposure of the animal to determinants of the toleragen. As the determinants composing the toleragen are generally a portion of those antigenic determinants comprising the immunogen (i.e. the background epitopes), the reduced antibody response to the background epitopes upon challenge with the immunogen can act to unmask the antibody response to the immunorecessive epitopes of the immunogen. By unmasked, it is meant that the population of antibody-producing cells directed to the immunorecessive epitopes effectively becomes a greater percentage of the overall population of antibody-producing cells in the animal (see Williams et al. (1992) *Biotechniques* 12:842-847).

In the subject method most preferred, immunotolerizing means includes subtractive immunization for enriching a pool of B-cells for clones producing antibodies specific for rare epitopes. Generally, subtractive immunization is a two-step procedure. Step one is a suppression step in which a state of tolerance is induced in the immune system of a host animal to a specific set of molecules, the tolerogen. Step two is an immunizing step in which another set of molecules, the immunogen, is introduced to the immune system. The molecules comprising the tolerogen are generally a subset of those comprising the immunogen. Ideally, the only molecules to which the immune system will generate the antibodies after exposure to the immunogen are those molecules present in the immunogen but not present in the tolerogen. Two main approaches have been used for subtractive immunization: neonatal tolerization and chemical immunosuppression.

In one embodiment of the invention, neonatal tolerization is utilized to generate an enriched pool of B-cells. Neonatal tolerization utilizes the self-tolerization process of the developing immune system. For each species, a discrete developmental period exists during which the immune system classifies all molecules present in the body as self, resulting in an induced state of immunological tolerance to those molecules (Billingham et al. (1953) *Nature* 172:603-606; Golumbeski et al. (1986) *Anal Biochem* 154:373-381; Hasek et al. (1979) *Immunol Rev* 46:3-26; Reading (1982) *J Immunol Methods* 53:261-291; and Streilen et al. (1979) *Immunol Rev* 46:125-146). Subsequent exposure to any molecules present during this stage will be met with immunological unresponsiveness. For subtractive immunization, mice (or other host animals) are neonatally exposed to the tolerogen. When these animals are immunologically mature, they are exposed to the immunogen. Theoretically, the immune system should be immunologically responsive only to those molecules in the immunogen, but not in the tolerogen.

In another embodiment of the subject method, chemical immunosuppression is the immunotolerizing means employed to generate an enriched B-cell population for subsequent cloning of variable region genes (V-genes). For example, chemical immunosuppression via the cytotoxic drug cyclophosaphamide is technique useful for subtractive immunization (Ahmed et al. (1984) *J Am Acad Dermatol* 11:1115-1126; Matthew et al. (1983) *CSH Symp*

Quant Biol 48:625-631; Matthew et al. (1987) *J Immunol Methods* 100:73-82; and Turk et al. (1972) *Immunology* 23:493-501). Application of the chemical cyclophosphamide to animals exposed to a foreign antigen selectively kills B-cells that have been stimulated to proliferate in response to the presence of the foreign antigenic molecules. After cyclophosphamide treatment, subsequent exposure to those molecules results in a reduced immunological response. As a subtractive immunization technique, animals are first exposed to the foreign antigenic molecule (i.e. the tolerogen), and are then injected with cyclophosphamide. After the drug has been allowed to clear, the animals are exposed to the immunogen. Theoretically, the immune system should be immunologically responsive only to those epitopes of the immunogen that are not found in the tolerogen.

Other subtractive immunization protocols are also available for use in the subject method, and include, for example, the use of interleukin-targeted toxins. For instance, IL-2-toxin fusion proteins (Kelley et al. (1988) *PNAS* 85:3980-3984) and IL-4-toxin fusion proteins (Lakkis et al. (1991) *Eur J Immunol* 21:2253-2258) can be used to selectively induce tolerance to the epitopes of a toleragen.

II. Generating Antibody Gene Libraries

After application of an immunotolerization step, the antibody repertoire of the resulting B-cell pool is cloned. Methods are generally known, and can be applied in the subject method, for directly obtaining the DNA sequence of the variable regions of a diverse population of immunoglobulin molecules by using a mixture of oligomer primers and PCR. For instance, mixed oligonucleotide primers corresponding to the 5' leader (signal peptide) sequences and/or framework 1 (FR1) sequences, as well as primer to a conserved 3' constant region primer can be used for PCR amplification of the heavy and light chain variable regions from a number of murine antibodies (Larrick et al. (1991) Biotechniques 11: 152-156). A similar strategy can also been used to amplify human heavy and light chain variable regions from human antibodies (Larrick et al. (1991) Methods: Companion to Methods in Enzymology 2: 106-110). The ability to clone human immunoglobulin V-genes takes on special significance in light of advancements in creating human antibody repertoires in transgenic animals (see, for example, Bruggeman et al. (1993) *Year Immunol* 7:33-40; Tuaillon et al. (1993) *PNAS* 90:3720-3724; Bruggeman et al. (1991) *Eur J Immunol* 21:1323-1326; and Wood et al. PCT publication WO 91/00906).

In an illustrative embodiment, RNA is isolated from mature B cells of, for example, peripheral blood cells, bone marrow, or spleen preparations, using standard protocols (e.g., U.S. Patent No. 4,683,202; Orlandi, et al. *PNAS* (1989) 86:3833-3837; Sastry et al., *PNAS* (1989) 86:5728-5732; and Huse et al. (1989) *Science* 246:1275-1281.) First-strand cDNA is

synthesized using primers specific for the constant region of the heavy chain(s) and each of the κ and λ light chains, as well as primers for the signal sequence. Using variable region PCR primers, such as those shown in Figures 1A and 1B (for mouse) or Figure 6 (for human), the variable regions of both heavy and light chains are amplified, each alone or in combination, and ligated into appropriate vectors for further manipulation in generating the display packages.

Oligonucleotide primers useful in amplification protocols may be unique or degenerate or incorporate inosine at degenerate positions. Restriction endonuclease recognition sequences may also be incorporated into the primers to allow for the cloning of the amplified fragment into a vector in a predetermined reading frame for expression.

III. Variegated Antibody Display

The V-gene library cloned from the immunotolerance-derived antibody repertoire can be expressed by a population of display packages to form an antibody display library. With respect to the display package on which the variegated antibody library is manifest, it will be appreciated from the discussion provided herein that the display package will often preferably be able to be (i) genetically altered to encode at least a variable region of an antibody, (ii) maintained and amplified in culture, (iii) manipulated to display the antibody gene product in a manner permitting the antibody to interact with a target epitope during an affinity separation step, and (iv) affinity separated while retaining the antibody gene such that the sequence of the antibody gene can be obtained. In preferred embodiments, the display remains viable after affinity separation.

Ideally, the display package comprises a system that allows the sampling of very large variegated antibody display libraries, rapid sorting after each affinity separation round, and easy isolation of the antibody gene from purified display packages. The most attractive candidates for this type of screening are prokaryotic organisms and viruses, as they can be amplified quickly, they are relatively easy to manipulate, and large number of clones can be created. Preferred display packages include, for example, vegetative bacterial cells, bacterial spores, and most preferably, bacterial viruses (especially DNA viruses). However, the present invention also contemplates the use of eukaryotic cells (other than cells which naturally produce antibodies, i.e. B-cells), including yeast and their spores, as potential display packages.

In addition to commercially available kits for generating phage display libraries (e.g. the Pharmacia *Recombinant Phage Antibody System*, catalog no. 27-9400-01; and the Stratagene *SurfZAP*™ phage display kit, catalog no. 240612), examples of methods and

reagents particularly amenable for use in generating the variegated antibody display library of the present invention can be found in, for example, the Ladner et al. U.S. Patent No. 5,223,409; the Kang et al. International Publication No. WO 92/18619; the Dower et al. International Publication No. WO 91/17271; the Winter et al. International Publication WO 92/20791; the Markland et al. International Publication No. WO 92/15679; the Breitling et al. International Publication WO 93/01288; the McCafferty et al. International Publication No. WO 92/01047; the Garrard et al. International Publication No. WO 92/09690; the Ladner et al. International Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum Antibod Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffths et al. (1993) *EMBO J* 12:725-734; Hawkins et al. (1992) *J Mol Biol* 226:889-896; Clackson et al. (1991) *Nature* 352:624-628; Gram et al. (1992) *PNAS* 89:3576-3580; Garrad et al. (1991) *Bio/Technology* 9:1373-1377; Hoogenboom et al. (1991) *Nuc Acid Res* 19:4133-4137; and Barbas et al. (1991) *PNAS* 88:7978-7982.

When the display is based on a bacterial cell, or a phage which is assembled periplasmically, the display means of the package will comprise at least two components. The first component is a secretion signal which directs the recombinant antibody to be localized on the extracellular side of the cell membrane (of the host cell when the display package is a phage). This secretion signal is characteristically cleaved off by a signal peptidase to yield a processed, "mature" antibody. The second component is a display anchor protein which directs the display package to associate the antibody with its outer surface. As described below, this anchor protein can be derived from a surface or coat protein native to the genetic package.

When the display package is a bacterial spore, or a phage whose protein coating is assembled intracellularly, a secretion signal directing the antibody to the inner membrane of the host cell is unnecessary. In these cases, the means for arraying the variegated antibody library comprises a derivative of a spore or phage coat protein amenable for use as a fusion protein.

The antibody component of the display will comprise, at a minimum, one of either the V_H or V_L regions cloned from B cells isolated in the subtractive immunization step. It will be appreciated, however, that the V_H regions and/or the V_L regions may contain, in addition to the variable portion of the antibodies, all or a portion of the constant regions. Typically, the display library will include variable regions of both heavy and light chains in order to generate at least an Fv fragment. For clarity, embodiments described herein detail the minimal antibody display as comprising the use of cloned V_H regions to construct the fusion protein with the display anchor protein. However, it should be readily understood that similar embodiments are possible in which the role of the V_L and V_H chains are reversed in the construction of the display library.

Under certain circumstances, the V_H portion of the antibody display is derived from isolated cells of the subtractive immunization step, but the V_L chain is either absent or is a "fixed" V_L (i.e. the same V_L chain for every antibody of the display). Where, for example, the V_L portion of the display is fixed, the V_L chain can be contributed from a gene construct other than the construct encoding the V_H chain, or from the host cell itself (i.e. a light chain producing myeloma cell), or added exogenously to the packages so as to recombine with V_H chains already displayed on their surface. However, it will generally be preferred that the V_L chain is derived from a variegated V_L library also cloned from the same population of B cells from which the V_H gene is cloned, in which case a preferred embodiment places the V_L gene in the same construct as the V_H gene such that both may be readily recovered together.

When the desired antibody display is a multi-chain antibody (e.g. V_H and V_L are separate polypeptide chains), the cDNA encoding the light chain may be cloned directly into an appropriate site of the vector containing the heavy chain-coat protein library; or, alternatively, the light chain may be cloned as a separate library in a different plasmid vector, amplified, and subsequently the fragments cloned into the vector library encoding the heavy chain. In such circumstances, the V_L chain is cloned so that it is expressed with a signal peptide leader sequence that will direct its secretion into the periplasm of the host cell. For example, several leader sequences have been shown to direct the secretion of antibody sequences in E. coli, such as OmpA (Hsiung et al. *Bio/Technology* (1986) 4:991-995), and (Better et al. *Science* 240:1041-1043), phoA (Skerra and Pluckthun, *Science* (1988) 240:1038).

In the instance wherein the display package is a phage, the cloning site for the V_L chain sequences in the phagemid should be placed so that it does not substantially interfere with normal phage function. One such locus is the intergenic region as described by Zinder and Boeke, (1982) *Gene* 19:1-10. In an illustrative embodiment comprising an M13 phage display library, the V_L sequence is preferably expressed at an equal or higher-level than the H_L-cpIII product (described below) to maintain a sufficiently high V_L concentration in the periplasm and provide efficient assembly (association) of V_L with V_H chains. For instance, a phagemid can be constructed to encode, as separate genes, both a V_H/coat fusion protein and a V_L chain. Under the appropriate induction, both chains are expressed and allowed to assemble in the periplasmic space of the host cell, the assembled antibody being linked to the phage particle by virtue of the V_H chain being a portion of a coat protein fusion construct.

The number of possible combinations of heavy and light chains probably exceeds 10^{12}. To sample as many combinations as possible depends, in part, on the ability to recover large numbers of transformants. For phage with plasmid-like forms (as filamentous phage), electrotransformation provides an efficiency comparable to that of phage-transfection with *in vitro* packaging, in addition to a very high capacity for DNA input. This allows large amounts

of vector DNA to be used to obtain very large numbers of transformants. The method described by Dower et al. (1988) *Nucleic Acids Res.*, 16:6127-6145, for example, may be used to transform fd-tet derived recombinants at the rate of about 10^7 transformants/ug of ligated vector into E. coli (such as strain MC1061), and libraries may be constructed in fd-tet Bl of up to about 3 x 10^8 members or more. Increasing DNA input and making modifications to the cloning protocol within the ability of the skilled artisan may produce increases of greater than about 10- fold in the recovery of transformants, providing libraries of up to 10^{10} or more recombinants.

In other embodiments, the V region domains of heavy and light chains can be expressed on the same polypeptide, joined by a flexible linker to form a single-chain Fv fragment, and the scFV gene subsequently cloned into the desired expression vector or phage genome. As generally described in McCafferty et al., *Nature* (1990) 348:552-554, complete V_H and V_L domains of an antibody, joined by a flexible $(Gly_4\text{-}Ser)_3$ linker can be used to produce a single chain antibody which can render the display package separable based on antigen affinity.

As will be apparent to those skilled in the art, in embodiments wherein high affinity antibodies are sought, an important criteria for the present selection method can be that it is able to discriminate between antibodies of different affinity for a particular antigen, and preferentially enrich for the antibodies of highest affinity. Applying the well known principles of antibody affinity and valence (i.e. avidity), it is understood that manipulating the display package to be rendered effectively monovalent can allow affinity enrichment to be carried out for generally higher binding affinities (i.e. binding constants in the range of 10^6 to 10^{10} M^{-1}) as compared to the broader range of affinities isolable using a multivalent display package. To generate the monovalent display, the natural (i.e. wild-type) form of the surface or coat protein used to anchor the antibody to the display can be added at a high enough level that it almost entirely eliminates inclusion of the antibody fusion protein in the display package. Thus, a vast majority of the display packages can be generated to include no more than one copy of the antibody fusion protein (see, for example, Garrad et al. (1991) *Bio/Technology* 9:1373-1377). In a preferred embodiment of a monovalent display library, the library of display packages will comprise no more than 5 to 10% polyvalent displays, and more preferably no more than 2% of the display will be polyvalent , and most preferably, no more than 1% polyvalent display packages in the population. The source of the wild-type anchor protein can be, for example, provided by a copy of the wild-type gene present on the same construct as the antibody fusion protein, or provided by a separate construct altogether. However, it will be equally clear that by similar manipulation, polyvalent displays can be generated to isolate a broader range of binding affinities. Such antibodies can be useful, for example, in purification protocols where avidity can be desirable.

i) Phages As Display Packages

Bacteriophage are attractive prokaryotic-related organisms for use in the subject method. Bacteriophage are excellent candidates for providing a display system of the variegated antibody library as there is little or no enzymatic activity associated with intact mature phage, and because their genes are inactive outside a bacterial host, rendering the mature phage particles metabolically inert. In general, the phage surface is a relatively simple structure. Phage can be grown easily in large numbers, they are amenable to the practical handling involved in many potential mass screening programs, and they carry genetic information for their own synthesis within a small, simple package. As the antibody gene is inserted into the phage genome, choosing the appropriate phage to be employed in the subject method will generally depend most on whether (i) the genome of the phage allows introduction of the antibody gene either by tolerating additional genetic material or by having replaceable genetic material; (ii) the virion is capable of packaging the genome after accepting the insertion or substitution of genetic material; and (iii) the display of the antibody on the phage surface does not disrupt virion structure sufficiently to interfere with phage propagation.

One concern presented with the use of phage is that the morphogenetic pathway of the phage determines the environment in which the antibody will have opportunity to fold. Periplasmically assembled phage are preferred as the displayed antibodies will generally contain essential disulfides, and such antibodies may not fold correctly within a cell. However, in certain embodiments in which the display package forms intracellularly (e.g., where λ phage are used), it has been demonstrated that the antibody may assume proper folding after the phage is released from the cell.

Another concern related to the use of phage, but also pertinent to the use of bacterial cells and spores as well, is that multiple infections could generate hybrid displays that carry the gene for one particular antibody yet have at least one or more different antibodies on their surfaces. Therefore, it can be preferable, though optional,to minimize this possibility by infecting cells with phage under conditions resulting in a low multiple-infection. However, there may be circumstances in which high multiple-infection conditions would be desirable, such as to increase homologous recombination events between gene constructs encoding the antibody display in order to further expand the repertoire of the antibody display library.

For a given bacteriophage, the preferred display means is a protein that is present on the phage surface (e.g. a coat protein). Filamentous phage can be described by a helical lattice; isometric phage, by an icosahedral lattice. Each monomer of each major coat protein sits on a lattice point and makes defined interactions with each of its neighbors. Proteins that fit into the lattice by making some, but not all, of the normal lattice contacts are likely to

destabilize the virion by aborting formation of the virion as well as by leaving gaps in the virion so that the nucleic acid is not protected. Thus in bacteriophage, unlike the cases of bacteria and spores, it is generally important to retain in the antibody fusion proteins those residues of the coat protein that interact with other proteins in the virion. For example, when using the M13 cpVIII protein, the entire mature protein will generally be retained with the antibody fragment being added to the N-terminus of cpVIII, while on the other hand it can suffice to retain only the last 100 carboxy terminal residues (or even fewer) of the M13 cpIII coat protein in the antibody fusion protein.

Under the appropriate induction, the antibody library is expressed and allowed to assemble in the bacterial cytoplasm, such as when the λ phage is employed. The induction of the protein(s) may be delayed until some replication of the phage genome, synthesis of some of the phage structural-proteins, and assembly of some phage particles has occurred. The assembled protein chains then interact with the phage particles via the binding of the anchor protein on the outer surface of the phage particle. The cells are lysed and the phage bearing the library-encoded receptor protein (that corresponds to the specific library sequences carried in the DNA of that phage) are released and isolated from the bacterial debris.

To enrich for and isolate phage which contain cloned library sequences that encode a desired protein, and thus to ultimately isolate the nucleic acid sequences themselves, phage harvested from the bacterial debris are affinity purified. As described below, when an antibody which specifically binds a particular antigen or antigenic determinant is desired, the antigen or determinant can be used to retrieve phage displaying the desired antibody. The phage so obtained may then be amplified by infecting into host cells. Additional rounds of affinity enrichment followed by amplification may be employed until the desired level of enrichment is reached.

The enriched antibody-phage can also be screened with additional detection-techniques such as expression plaque (or colony) lift (see, e.g., Young and Davis, *Science* (1983) 222:778-782) whereby a labeled antigen is used as a probe. The phage obtained from the screening protocol are infected into cells, propagated, and the phage DNA isolated and sequenced, and/or recloned into a vector intended for gene expression in prokaryotes or eukaryotes to obtain larger amounts of the particular antibody selected.

In yet another embodiment, the antibody is also transported to an extra-cytoplasmic compartment of the host cell, such as the bacterial periplasm, but as a fusion protein with a viral coat protein. In this embodiment the desired protein (or one of its polypeptide chains if it is a multichain antibody) is expressed fused to a viral coat protein which is processed and transported to the cell inner membrane. Other chains, if present, are expressed with a secretion leader and thus are also transported to the periplasm or other intracellular by extra-

cytoplasmic location. The chains (e.g. heavy and light chains) present in the extra-cytoplasm then assemble into a complete antibody (or binding fragment thereof), The assembled molecules become incorporated into the phage by virtue of their attachment to the phage coat protein as the phage extrude through the host membrane and the coat proteins assemble around the phage DNA. The phage bearing the antibody complex may then be screened by affinity enrichment as described below.

a) Filamentous Phage

Filamentous bacteriophages, which include M13, fl, fd, Ifl, Ike, Xf, Pfl, and Pf3, are a group of related viruses that infect bacteria. They are termed filamentous because they are long, thin particles comprised of an elongated capsule that envelopes the deoxyribonucleic acid (DNA) that forms the bacteriophage genome. The F pili filamentous bacteriophage (Ff phage) infect only gram-negative bacteria by specifically adsorbing to the tip of F pili, and include fd, fl and M13.

Compared to other bacteriophage, filamentous phage in general are attractive and M13 in particular is especially attractive because: (i) the 3-D structure of the virion is known; (ii) the processing of the coat protein is well understood; (iii) the genome is expandable; (iv) the genome is small; (v) the sequence of the genome is known; (vi) the virion is physically resistant to shear, heat, cold, urea, guanidinium chloride, low pH, and high salt; (vii) the phage is a sequencing vector so that sequencing is especially easy; (viii) antibiotic-resistance genes have been cloned into the genome with predictable results (Hines et al. (1980) *Gene* 11:207-218); (ix) it is easily cultured and stored, with no unusual or expensive media requirements for the infected cells, (x) it has a high burst size, each infected cell yielding 100 to 1000 M13 progeny after infection; and (xi) it is easily harvested and concentrated (Salivar et al. (1964) *Virology* 24: 359-371). The entire life cycle of the filamentous phage M13, a common cloning and sequencing vector, is well understood. The genetic structure of M13 is well known, including the complete sequence (Schaller et al. in *The Single-Stranded DNA Phages* eds. Denhardt et al. (NY: CSHL Press, 1978)), the identity and function of the ten genes, and the order of transcription and location of the promoters, as well as the physical structure of the virion (Smith et al. (1985) *Science* 228:1315-1317; Raschad et al. (1986) *Microbiol Dev* 50:401-427; Kuhn et al. (1987) *Science* 238:1413-1415; Zimmerman et al. (1982) *J Biol Chem* 257:6529-6536; and Banner et al. (1981) *Nature* 289:814-816). Because the genome is small (6423 bp), cassette mutagenesis is practical on RF M13 (*Current Protocols in Molecular Biology,* eds. Ausubel et al. (NY: John Wiley & Sons, 1991)), as is single-stranded oligonucleotide directed mutagenesis (Fritz et al. in *DNA Cloning*, ed by Glover (Oxford, UK: IRC Press, 1985)). M13 is a plasmid and transformation system in itself, and an ideal sequencing vector. M13 can be grown on Rec- strains of *E. coli*. The M13 genome is expandable (Messing et al. in *The Single-Stranded DNA Phages*, eds Denhardt et

al. (NY: CSHL Press, 1978) pages 449-453; and Fritz et al., *supra*) and M13 does not lyse cells. Extra genes can be inserted into M13 and will be maintained in the viral genome in a stable manner.

The mature capsule or Ff phage is comprised of a coat of five phage-encoded gene products: cpVIII, the major coat protein product of gene VIII that forms the bulk of the capsule; and four minor coat proteins, cpIII and cpIV at one end of the capsule and cpVII and cpIX at the other end of the capsule. The length of the capsule is formed by 2500 to 3000 copies of cpVIII in an ordered helix array that forms the characteristic filament structure. The gene III-encoded protein (cpIII) is typically present in 4 to 6 copies at one end of the capsule and serves as the receptor for binding of the phage to its bacterial host in the initial phase of infection. For detailed reviews of Ff phage structure, see Rasched et al., *Microbiol. Rev.*, 50:401-427 (1986); and Model et al.,in *The Bacteriophages, Volume 2*, R. Calendar, Ed., Plenum Press, pp. 375-456 (1988).

The phage particle assembly involves extrusion of the viral genome through the host cell's membrane. Prior to extrusion, the major coat protein cpVIII and the minor coat protein cpIII are synthesized and transported to the host cell's membrane. Both cpVIII and cpIII are anchored in the host cell membrane prior to their incorporation into the mature particle. In addition, the viral genome is produced and coated with cpV protein. During the extrusion process, cpV-coated genomic DNA is stripped of the cpV coat and simultaneously recoated with the mature coat proteins.

Both cpIII and cpVIII proteins include two domains that provide signals for assembly of the mature phage particle. The first domain is a secretion signal that directs the newly synthesized protein to the host cell membrane. The secretion signal is located at the amino terminus of the polypeptide and targets the polypeptide at least to the cell membrane. The second domain is a membrane anchor domain that provides signals for association with the host cell membrane and for association with the phage particle during assembly. This second signal for both cpVIII and cpIII comprises at least a hydrophobic region for spanning the membrane.

The 50 amino acid mature gene VIII coat protein (cpVIII) is synthesized as a 73 amino acid precoat (Ito et al. (1979) *PNAS* 76:1199-1203). cpVIII has been extensively studied as a model membrane protein because it can integrate into lipid bilayers such as the cell membrane in an asymmetric orientation with the acidic amino terminus toward the outside and the basic carboxy terminus toward the inside of the membrane. The first 23 amino acids constitute a typical signal-sequence which causes the nascent polypeptide to be inserted into the inner cell membrane. An E. coli signal peptidase (SP-I) recognizes amino acids 18, 21, and 23, and, to a lesser extent, residue 22, and cuts between residues 23 and 24

of the precoat (Kuhn et al. (1985) *J. Biol. Chem.* 260:15914-15918; and Kuhn et al. (1985) *J. Biol. Chem.* 260:15907-15913). After removal of the signal sequence, the amino terminus of the mature coat is located on the periplasmic side of the inner membrane; the carboxy terminus is on the cytoplasmic side. About 3000 copies of the mature coat protein associate side-by-side in the inner membrane.

The sequence of gene VIII is known, and the amino acid sequence can be encoded on a synthetic gene. Mature gene VIII protein makes up the sheath around the circular ssDNA. The gene VIII protein can be a suitable anchor protein because its location and orientation in the virion are known (Banner et al. (1981) *Nature* 289:814-816). Preferably, the antibody is attached to the amino terminus of the mature M13 coat protein to generate the phage display library. As set out above, manipulation of the concentration of both the wild-type cpVIII and Ab/cpVIII fusion in an infected cell can be utilized to decrease the avidity of the display and thereby enhance the detection of high affinity antibodies directed to the target epitope(s).

Another vehicle for displaying the antibody is by expressing it as a domain of a chimeric gene containing part or all of gene III. When monovalent displays are required, expressing the V-gene as a fusion protein with gpIII can be a preferred embodiment, as manipulation of the ratio of wild-type gpIII to chimeric gpIII during formation of the phage particles can be readily controlled. This gene encodes one of the minor coat proteins of M13. Genes VI, VII, and IX also encode minor coat proteins. Each of these minor proteins is present in about 5 copies per virion and is related to morphogenesis or infection. In contrast, the major coat protein is present in more than 2500 copies per virion. The gene VI, VII, and IX proteins are present at the ends of the virion; these three proteins are not post-translationally processed (Rasched et al. (1986) *Ann Rev. Microbiol.* 41:507-541). In particular, the single-stranded circular phage DNA associates with about five copies of the gene III protein and is then extruded through the patch of membrane-associated coat protein in such a way that the DNA is encased in a helical sheath of protein (Webster et al. in *The Single-Stranded DNA Phages,* eds Dressler et al. (NY:CSHL Press, 1978).

Manipulation of the sequence of cpIII has demonstrated that the C-terminal 23 amino acid residue stretch of hydrophobic amino acids normally responsible for a membrane anchor function can be altered in a variety of ways and retain the capacity to associate with membranes. Ff phage-based expression vectors were first described in which the cpIII amino acid residue sequence was modified by insertion of polypeptide "epitopes" (Parmely et al., *Gene* (1988) 73:305-318; and Cwirla et al., *PNAS* (1990) 87:6378-6382) or an amino acid residue sequence defining a single chain antibody domain (McCafferty et al., *Science* (1990) 348:552-554). It has been demonstrated that insertions into gene III can result in the production of novel protein domains on the virion outer surface. (Smith (1985) *Science* 228:1315-1317; and de la Cruz et al. (1988) *J. Biol. Chem.* 263:4318-4322). The antibody

gene may be fused to gene III at the site used by Smith and by de la Cruz et al., at a codon corresponding to another domain boundary or to a surface loop of the protein, or to the amino terminus of the mature protein.

Generally, the successful cloning strategy utilizing a phage coat protein, such as cpIII of filamentous phage fd, will provide: (1) expression of an antibody chain fused to the N-terminus of a coat protein (e.g., cpIII) and transport to the inner membrane of the host where the hydrophobic domain in the C-terminal region of the coat protein anchors the fusion protein in the membrane, with the N-terminus containing the antibody chain protruding into the periplasmic space and available for interaction with a second or subsequent chain (e.g., V_L to form an Fv or Fab fragment) which is thus attached to the coat protein; and (2) adequate expression of a second or subsequent polypeptide chain if present (e.g., V_L) and transport of this chain to the soluble compartment of the periplasm.

Similar constructions could be made with other filamentous phage. Pf3 is a well known filamentous phage that infects *Pseudomonas aerugenosa* cells that harbor an IncP-I plasmid. The entire genome has been sequenced ((Luiten et al. (1985) *J. Virol.* 56:268-276) and the genetic signals involved in replication and assembly are known (Luiten et al. (1987) DNA 6:129-137). The major coat protein of PF3 is unusual in having no signal peptide to direct its secretion. The sequence has charged residues ASP-7, ARG-37, LYS-40, and PHE44 which is consistent with the amino terminus being exposed. Thus, to cause an antibody to appear on the surface of Pf3, a tripartite gene can be constructed which comprises a signal sequence known to cause secretion in *P. aerugenosa*, fused in-frame to a gene fragment encoding the antibody sequence, which is fused in-frame to DNA encoding the mature Pf3 coat protein. Optionally, DNA encoding a flexible linker of one to 10 amino acids is introduced between the antibody gene fragment and the Pf3 coat-protein gene. This tripartite gene is introduced into Pf3 so that it does not interfere with expression of any Pf3 genes. Once the signal sequence is cleaved off, the antibody is in the periplasm and the mature coat protein acts as an anchor and phage-assembly signal.

b) Bacteriophage ϕX174

The bacteriophage ϕX174 is a very small icosahedral virus which has been thoroughly studied by genetics, biochemistry, and electron microscopy (see *The Single Stranded DNA Phages* (eds. Den hardt et al. (NY:CSHL Press, 1978)). Three gene products of ϕX174 are present on the outside of the mature virion: F (capsid), G (major spike protein, 60 copies per virion), and H (minor spike protein, 12 copies per virion). The G protein comprises 175 amino acids, while H comprises 328 amino acids. The F protein interacts with the single-stranded DNA of the virus. The proteins F, G, and H are translated from a single mRNA in the viral infected cells. As the virus is so tightly constrained because several of its

genes overlap, ϕX174 is not typically used as a cloning vector due to the fact that it can accept very little additional DNA. However, mutations in the viral G gene (encoding the G protein) can be rescued by a copy of the wild-type G gene carried on a plasmid that is expressed in the same host cell (Chambers et al. (1982) *Nuc Acid Res* 10:6465-6473). In one embodiment, one or more stop codons are introduced into the G gene so that no G protein is produced from the viral genome. The variegated antibody gene library can then be fused with the nucleic acid sequence of the H gene. An amount of the viral G gene equal to the size of antibody gene fragment is eliminated from the ϕX174 genome, such that the size of the genome is ultimately unchanged. Thus, in host cells also transformed with a second plasmid expressing the wild-type G protein, the production of viral particles from the mutant virus is rescued by the exogenous G protein source. Where it is desirable that only one antibody be displayed per ϕX174 particle, the second plasmid can further include one or more copies of the wild-type H protein gene so that a mix of H and Ab/H proteins will be predominated by the wild-type H upon incorporation into phage particles.

c) Large DNA Phage

Phage such as λ or T4 have much larger genomes than do M13 or ϕX174, and have more complicated 3-D capsid structures than Ml3 or ϕPX174, with more coat proteins to choose from. In embodiments of the invention whereby the antibody library is processed and assembled into a functional form and associates with the bacteriophage particles within the cytoplasm of the host cell, bacteriophage λ and derivatives thereof are examples of suitable vectors. The intracellular morphogenesis of phage λ can potentially prevent protein domains that ordinarily contain disulfide bonds from folding correctly. However, variegated libraries expressing a population of functional antibodies, including both heavy and light chain variable regions, have been generated in λ phage. (Huse et al. (1989) *Science* 246:1275-1281; Mullinax et al. (1990) *PNAS* 87:8095-8099; and Pearson et al. (1991) *PNAS* 88:2432-2436). Such strategies take advantage of the rapid construction and efficient transformation abilities of λ phage.

When used for expression of antibody sequences, such as V_H, V_L, Fv (variable region fragment) or Fab, library DNA may be readily inserted into a λ vector. For instance, variegated antibody libraries have been constructed by modification of λ ZAP II (Short et al. (1988) *Nuc Acid Res* 16:7583) comprising inserting both cloned heavy and light chain variable regions into the multiple cloning site of a λ ZAP II vector (Huse et al. *supra.*). To illustrate, a pair of λ vectors may be designed to be asymmetric with respect to restriction sites that flank the cloning and expression sequences. This asymmetry allows efficient recombination of libraries coding for separate chains of the active protein. Thus, a library expressing antibody light chain variable regions (V_L) may be combined with one expressing antibody heavy chain variable regions (V_H), thereby constructing combinatorial antibody or

Fab expression libraries. For instance, one λ vector is designed to serve as a cloning vector for antibody light chain sequences, and another λ vector is designed to serve as a cloning vector for antibody heavy chain sequences in the initial steps of library construction. A combinatorial library is constructed from the two λ libraries by crossing them at an appropriate restriction site. DNA is first purified from each library, and the right and left arms of each respective λ vector cleaved so as to leave the antibody chain sequences intact. The DNAs are then mixed and ligated, and only clones that result from proper assembly of reciprocal vectors reconstitute as viable phage (Huse et all, *supra.*)

ii) Bacterial Cells as Display Packages

Recombinant antibodies are able to cross bacterial membranes after the addition of bacterial leader sequences to the N-terminus of the protein (Better et al (1988) *Science* 240:1041-1043; and Skerra et al. (1988) *Science* 240:1038-1041). In addition, recombinant antibodies have been fused to outer membrane proteins for surface presentation. For example, one strategy for displaying antibodies on bacterial cells comprises generating a fusion protein by inserting the antibody into cell surface exposed portions of an integral outer membrane protein (Fuchs et al. (1991) *Bio/Technology* 9:1370-1372). In selecting a bacterial cell to serve as the display package, any well-characterized bacterial strain will typically be suitable, provided the bacteria may be grown in culture, engineered to display the antibody library on its surface, and is compatible with the particular affinity selection process practiced in the subject method. Among bacterial cells, the preferred display systems include *Salmonella typhirnurium*, *Bacillus subtilis*, *Pseudomonas aeruginosa*, *Vibrio cholerae*, *Klebsiella pneumonia*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Bacteroides nodosus*, *Moraxella bovis*, and especially *Escherichia coli*. Many bacterial cell surface proteins useful in the present invention have been characterized, and works on the localization of these proteins and the methods of determining their structure include Benz et al. (1988) *Ann Rev Microbiol* 42: 359-393; Balduyck et al. (1985) *Biol Chem Hoppe-Seyler* 366:9-14; Ehrmann et al (1990) *PNAS* 87:7574-7578; Heijne et al. (1990) *Protein Engineering* 4:109-112; Ladner et al. U.S. Patent No. 5,223,409; Ladner et al. WO88/06630; Fuchs et al. (1991) *Bio/technology* 9:1370-1372; and Goward et al. (1992) *TIBS* 18:136-140.

To further illustrate, the LamB protein of *E coli* is a well understood surface protein that can be used to generate a variegated library of antibodies on the surface of a bacterial cell (see, for example, Ronco et al. (1990) *Biochemie* 72:183-189; van der Weit et al. (1990) *Vaccine* 8:269-277; Charabit et al. (1988) *Gene* 70:181-189; and Ladner U.S. Patent No. 5,222,409). LamB of *E. coli* is a porin for maltose and maltodextrin transport, and serves as the receptor for adsorption of bacteriophages λ and K10. LamB is transported to the outer membrane if a functional N-terminal signal sequence is present (Benson et al. (1984) *PNAS* 81:3830-3834). As with other cell surface proteins, LamB is synthesized with a typical

signal-sequence which is subsequently removed. Thus, the variegated antibody gene library can be cloned into the LamB gene such that the resulting library of fusion proteins comprise a portion of LamB sufficient to anchor the protein to the cell membrane with the antibody fragment oriented on the extracellular side of the membrane. Secretion of the extracellular portion of the fusion protein can be facilitated by inclusion of the LamB signal sequence, or other suitable signal sequence, as the N-terminus of the protein.

The E. coli LamB has also been expressed in functional form in *S. typhimurium* (Harkki et al. (1987) *Mol Gen Genet* 209:607-611), *V. cholerae* (Harkki et al. (1986) *Microb Pathol* 1:283-288), and *K. pneumonia* (Wehmeier et al. (1989) *Mol Gen Genet* 215:529-536), so that one could display a population of antibodies in any of these species as a fusion to E. coli LamB. Moreover, *K. pneumonia* expresses a maltoporin similar to LamB which could also be used. In *P. aeruginosa*, the Dl protein (a homologue of LamB) can be used (Trias et al. (1988) *Biochem Biophys Acta* 938:493-496). Similarly, other bacterial surface proteins, such as PAL, OmpA, OmpC, OmpF, PhoE, pilin, BtuB, FepA, FhuA, IutA, FecA and FhuE, may be used in place of LamB as a portion of the display means in a bacterial cell.

iii) Bacterial Spores as Display Packages

Bacterial spores also have desirable properties as display package candidates in the subject method. For example, spores are much more resistant than vegetative bacterial cells or phage to chemical and physical agents, and hence permit the use of a great variety of affinity selection conditions. Also, Bacillus spores neither actively metabolize nor alter the proteins on their surface. However, spores have the disadvantage that the molecular mechanisms that trigger sporulation are less well worked out than is the formation of M13 or the export of protein to the outer membrane of *E. coli*, though such a limitation is not a serious detractant from their use in the present invention.

Bacteria of the genus Bacillus form endospores that are extremely resistant to damage by heat, radiation, desiccation, and toxic chemicals (reviewed by Losick et al. (1986) *Ann Rev Genet* 20:625-669). This phenomenon is attributed to extensive intermolecular cross-linking of the coat proteins. In certain embodiments of the subject method, such as those which include relatively harsh affinity separation steps, Bacillus spores can be the preferred display package. Endospores from the genus Bacillus are more stable than are, for example, exospores from Streptomyces. Moreover, *Bacillus subtilis* forms spores in 4 to 6 hours, whereas Streptomyces species may require days or weeks to sporulate. In addition, genetic knowledge and manipulation is much more developed for *B. subtilis* than for other spore-forming bacteria.

Viable spores that differ only slightly from wild-type are produced in *B. subtilis* even if any one of four coat proteins is missing (Donovan et al. (1987) *J Mol Biol* 196:1-10).

Moreover, plasmid DNA is commonly included in spores, and plasmid encoded proteins have been observed on the surface of Bacillus spores (Debro et al. (1986) *J Bacteriol* 165:258-268). Thus, it can be possible during sporulation to express a gene encoding a chimeric coat protein comprising an antibody of the variegated gene library, without interfering materially with spore formation.

To illustrate, several polypeptide components of *B. subtilis* spore coat (Donovan et al. (1987) *J Mol Biol* 196:1-10) have been characterized. The sequences of two complete coat proteins and amino-terminal fragments of two others have been determined. Fusion of the antibody sequence to cotC or cotD fragments is likely to cause the antibody to appear on the spore surface. The genes of each of these spore coat proteins are preferred as neither cotC or cotD are post-translationally modified (see Lader et al. U.S. Patent No. 5,223,409).

IV. Selecting Antibodies to a Target Antigen

Upon expression, the variegated antibody display is subjected to affinity enrichment in order to select for antibodies which bind preselected antigens. The term "affinity separation" or "affinity enrichment" includes, but is not limited to (1) affinity chromatography utilizing immobilizing antigens, (2) immunoprecipitation using soluble antigens, (3) fluorescence activated cell sorting, (4) agglutination, and (5) plaque lifts. In each embodiment, the library of display packages are ultimately separated based on the ability of the associated antibody to bind an epitope on the antigen of interest. See, for example, the Ladner et al. U.S. Patent No. 5,223,409; the Kang et al. International Publication No. WO 92/18619; the Dower et al. International Publication No. WO 91/17271; the Winter et al. International Publication WO 92/20791; the Markland et al. International Publication No. WO 92/15679; the Breitling et al. International Publication WO 93/01288; the McCafferty et al. International Publication No. WO 92/01047; the Garrard et al. International Publication No. WO 92/09690; and the Ladner et al. International Publication No. WO 90/02809. In most preferred embodiments, the display library will be pre-enriched for antibodies specific for the rare epitope by first contacting the display library with a source of the background epitope, such as the toleragen, in order to further remove antibodies which bind the background epitopes. Subsequently, the display package is contacted with the target antigen and antibodies of the display which are able to specifically bind the antigen are isolated.

With respect to affinity chromatography, it will be generally understood by those skilled in the art that a great number of chromatography techniques can be adapted for use in the present invention, ranging from column chromatography to batch elution, and including ELISA and biopanning techniques. Typically the target antigen is immobilized on an

insoluble carrier, such as sepharose or polyacrylamide beads, or, alternatively, the wells of a microtitre plate. As described below, in instances where no purified source of the target antigen is readily available, such as the case with many cell-specific markers, the cells on which the antigen is displayed may serve as the insoluble matrix carrier.

The population of display packages is applied to the affinity matrix under conditions compatible with the binding of the antibody to a target antigen. The population is then fractionated by washing with a solute that does not greatly effect specific binding of antibodies to the target antigen, but which substantially disrupts any non-specific binding of the display package to the antigen or matrix. A certain degree of control can be exerted over the binding characteristics of the antibodies recovered from the display library by adjusting the conditions of the binding incubation and subsequent washing. The temperature, pH, ionic strength, divalent cation concentration, and the volume and duration of the washing can select for antibodies within a particular range of affinity and specificity. Selection based on slow dissociation rate, which is usually predictive of high affinity, is a very practical route. This may be done either by continued incubation in the presence of a saturating amount of free hapten (if available), or by increasing the volume, number, and length of the washes. In each case, the rebinding of dissociated antibody-display package is prevented, and with increasing time, antibody-display packages of higher and higher affinity are recovered. Moreover, additional modifications of the binding and washing procedures may be applied to find antibodies with special characteristics. The affinities of some antibodies are dependent on ionic strength or cation concentration. This is a useful characteristic for antibodies to be used in affinity purification of various proteins when gentle conditions for removing the protein from the antibody are required. Specific examples are antibodies which depend on Ca^{++} for binding activity and which released their haptens in the presence of EGTA. (see, Hopp et al. (1988) *Biotechnology* 6:1204-1210). Such antibodies may be identified in the recombinant antibody library by a double screening technique isolating first those that bind hapten in the presence of Ca^{++}, and by subsequently identifying those in this group that fail to bind in the presence of EGTA.

After "washing" to remove non-specifically bound display packages, when desired, specifically bound display packages can be eluted by either specific desorption (using excess antigen) or non-specific desorption (using pH, polarity reducing agents, or chaotropic agents). In preferred embodiments, the elution protocol does not kill the organism used as the display package such that the enriched population of display packages can be further amplified by reproduction. The list of potential eluants includes salts (such as those in which one of the counter ions is Na^+, NH_4^+, Rb^+, SO_4^{2-}, $H_2PO_4^-$, citrate, K^+, Li^+, Cs^+, HSO_4^-, CO_3^{2-}, Ca^{2+}, Sr^{2+}, Cl^-, PO_4^{2-}, HCO_3^-, Mg_2^+, Ba_2^+, Br^-, HPO_4^{2-}, or acetate), acid, heat, and, when available, soluble forms of the target antigen (or analogs thereof). Because bacteria

continue to metabolize during the affinity separation step and are generally more susceptible to damage by harsh conditions, the choice of buffer components (especially eluates) can be more restricted when the display package is a bacteria rather than for phage or spores. Neutral solutes, such as ethanol, acetone, ether, or urea, are examples of other agents useful for eluting the bound display packages.

In preferred embodiments, affinity enriched display packages are iteratively amplified and subjected to further rounds of affinity separation until enrichment of the desired binding activity is detected. In certain embodiments, the specifically bound display packages, especially bacterial cells, need not be eluted per se, but rather, the matrix bound display packages can be used directly to inoculate a suitable growth media for amplification.

Where the display package is a phage particle, the fusion protein generated with the coat protein can interfere substantially with the subsequent amplification of eluted phage particles, particularly in embodiments wherein the cpIII protein is used as the display anchor. Even though present in only one of the 5-6 tail fibers, some antibody constructs because of their size and/or sequence, may cause severe defects in the infectivity of their carrier phage. This causes a loss of phage from the population during reinfection and amplification following each cycle of panning. In one embodiment, the antibody can be derived on the surface of the display package so as to be susceptible to proteolytic cleavage which severs the covalent linkage of at least the antigen binding sites of the displayed antibody from the remaining package. For instance, where the cpIII coat protein of M13 is employed, such a strategy can be used to obtain infectious phage by treatment with an enzyme which cleaves between the antibody portion and cpIII portion of a tail fiber fusion protein (e.g. such as the use of an enterokinase cleavage recognition sequence).

To further minimize problems associated with defective infectivity, DNA prepared from the eluted phage can be transformed into host cells by electroporation or well known chemical means. The cells are cultivated for a period of time sufficient for marker expression, and selection is applied as typically done for DNA transformation. The colonies are amplified, and phage harvested for a subsequent round(s) of panning.

After isolation of display packages which encode antibodies having a desired binding specificity for the immunorecessive epitope, the nucleic acid encoding the V-genes for each of the purified display packages can be recloned in a suitable eukaryotic or prokaryotic expression vector and transfected into an appropriate host for production of large amounts of protein. Where, for example, the isolated V-gene lacks a portion of a constant region and it is desirable that the missing portion be provided, simple molecular cloning techniques can be used to add back the missing portions. The binding affinity of the antibody can be confirmed by well known immunoassay techniques with the target epitope (see, for example, Harlow

and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988)).

V. Further Manipulation of Antibodies, Antibody Compositions, and Immunoassay Kits

Another aspect of the present invention concerns chimeric antibodies, e.g., altered antibodies in which at least the antigen binding portion of an immunoglobulin isolated by the method described above is cloned into another protein, preferably another antibody. Among embodiments of chimeric antibodies contemplated by the present invention, further manipulation of the subject antibodies can be used to complete the portion of the constant region isolated from the V-gene library, as well as to facilitate "class switching" whereby all or a portion of the constant region of the antibody isolated from the V-gene library is replaced with a different constant region, e.g., with the constant region(s) from a different IgG, such as IgG1, IgG2 or IgG3, or the constant region(s) from one of IgE, IgA, IgD or IgM. In similar fashion, single chain antibodies and other recombinant fragments can be generated from the cloned genes.

When antibodies produced in non-human subjects are used therapeutically in humans, they are recognized to varying degrees as foreign and an immune response may be generated in the patient. Accordingly, recombinant manipulation of the isolated antibody gene, where derived from a non-human V-gene library such as described in the Examples below, can be used to "humanize" the antibody. The term "humanized antibody" is used to describe a molecule having an antigen binding site derived from an immunoglobulin from a non-human species, the remaining immunoglobulin-derived portions of the molecule, as necessary to substantially reduce the immunogenicity of the molecule in human subjects, being derived from a human immunoglobulin. In a humanized antibody, the antigen binding site may include, for example, either complete variable domains fused to constant domains, or only the CDRs grafted to the appropriate framework regions in human variable domains. Such antibodies are the equivalents of the recombinant antibodies described above, but may be less immunogenic when administered to humans, and therefore more likely to be tolerated upon injected in a patient.

In an illustrative embodiment, any of the H3-3, FB3-2 or F4-7 antibodies described in the Examples below can be prepared to include human constant regions for each of the heavy and light chains of these mouse-derived genes. For example, the portion of the antibody gene encoding the murine constant region can be substituted with a gene encoding a human constant region (see Robinson et al., International Patent Publication PCT/US86/02269; Akira, et al., European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison et al., European Patent Application 173,494; Neuberger et

al., PCT Application WO 86/01533; Cabilly et al. U.S. Patent No. 4,816,567; Cabilly et al., European Patent Application 125,023; Better et al. (1988) *Science* 240:1041-1043; Liu et al. (1987) *PNAS* 84:3439-3443; Liu et al. (1987) *J. Immunol.* 139:3521-3526; Sun et al. (1987) *PNAS* 84:214-218; Nishimura et al. (1987) *Canc. Res.* 47:999-1005; Wood et al. (1985) *Nature* 314:446-449; and Shaw et al. (1988) *J. Natl Cancer Inst.* 80:1553-1559).

The subject antibodies can also be "humanized" by replacing portions of the variable region not involved in antigen binding with equivalent portions from human variable regions. General reviews of "humanized" chimeric antibodies are provided by Morrison, S. L. (1985) *Science* 229:1202-1207; and by Oi et al. (1986) *BioTechniques* 4:214. Those methods include isolating, manipulating, and expressing the nucleic acid sequences that encode all or part of an immunoglobulin variable region from at least one of a heavy or light chain. Sources of such nucleic acids are well known to those skilled in the art. The cDNA encoding the chimeric antibody, or fragment thereof, can then be cloned into an appropriate expression vector. Suitable "humanized" antibodies can be alternatively produced by CDR replacement (see U.S. Patent 5,225,539 to Winter; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan et al. (1988) *Science* 239:1534; and Beidler et al. (1988) *J. Immunol.* 141:4053-4060).

The DNA sequence encoding the chimeric variable domain may be prepared by oligonucleotide synthesis. This requires that at least the framework region sequence of the first antibody and at least the CDRs sequences of the subject antibody are known or can be readily determined. Determining these sequences, the synthesis of the DNA from oligonucleotides and the preparation of suitable vectors each involve the use of known techniques which can readily be carried out by a person skilled in the art in light of the teaching given herein.

Alternatively, the DNA sequence encoding the altered variable domain may be prepared by primer directed oligonucleotide site-directed mutagenesis. This technique in essence involves hybridizing an oligonucleotide coding for a desired mutation with a single strand of DNA containing the mutation and using the single strand as a template for extension of the oligonucleotide to produce a strand containing the mutation. This technique, in various forms, is described by: Zoller et al. (1982) *Nuc Acids Res* 10:6487-6500; Norriset al. (1983) *Nuc Acids Res* 11:5103-5112; Zoller et al. (1984) *DNA* 3:479-488; and Kramer et al. (1982) *Nuc Acids Res* 10:6475-6485. For various reasons, this technique in its simplest form does not always produce a high frequency of mutation. However, an improved technique for introducing both single and multiple mutations in an M13 based vector has been described by Carter et al. (1985) *Nuc Acids Res* 13:4431-4443. Using a long oligonucleotide, it has proved possible to introduce many changes simultaneously (e.g., see Carter et al., *supra*) and thus single oligonucleotides, each encoding a CDR, can be used to introduce the three CDRs from the subject antibody into the framework regions of a human antibody (see also U.S. Patent

5,345,847 to Liu et al.). Not only is this technique less laborious than total gene synthesis, but it represents a particularly convenient way of expressing a variable domain of required specificity, as it can be simpler than tailoring an entire V_H domain for insertion into an expression plasmid.

The oligonucleotides used for site-directed mutagenesis may be prepared by oligonucleotide synthesis or may be isolated from DNA coding for the variable domain of the subject antibody by use of suitable restriction enzymes. Such long oligonucleotides will generally be at least 30 residues long and may be up to or over 80 residues in length.

In yet another embodiment, PCR techniques for generating fusion proteins can be used to generate the chimeric antibody. PCR amplification of gene fragments, both CDR and FR regions, can be carried out using anchor primers which give rise to complementary overhangs between two consecutive CDR and FR fragments which can subsequently be annealed to generate a chimeric V-gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992).

The antigen binding sites of the subject antibodies can also be used to generate a fusion protein which includes protein sequences from non-immunoglobulin molecules. For example, such chimeric antibodies can include: proteins domains which render the protein cytotoxic or cytostatic, such as the addition of *Pseudomonas* exotoxin or *Diphtheria* toxin domains (see, for example, Jung et al. (1994) *Proteins* 19:35-47; Seetharam et al. (1991) *J Biol Chem* 266:17376-17381; and Nichols et al. (1993) *J Biol Chem* 268:5302-5308); DNA-binding polypeptides for facilitating DNA transport (see, for example, U.S. patent 5,166,320); catalytic domains which provide an enzymatic activity associated with the immunoglobulin, such as a phosphatase or peroxidase activity; and purification polypeptides to simplify purification of the antibody, such as a glutathione-S-transferase polypeptide for purification of the antibody with a glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. (N.Y.: John Wiley & Sons, 1991)), or a poly-(His)/enterokinase cleavage site sequence to permit purification of the poly(His)-antibody by affinity chromatography using a Ni^{2+} metal resin (e.g., see Hochuli et al. (1987) *J. Chromatography* 411:177; and Janknecht et al. *PNAS* 88:8972).

The present invention also makes available isolated forms of the subject antibodies which are isolated from, or otherwise substantially free of other cellular and extracellular proteins, especially antigenic proteins, or other extracellular factors, with which the antibodies normally bind. The term "substantially free of other cellular or extracellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of the subject antibodies having less than 20% (by dry weight) contaminating protein, and preferably having less than

5% contaminating protein. Functional forms of the subject antibodies can be prepared, for the first time, as purified preparations by using a cloned gene as described herein. By "purified", it is meant, when referring to a peptide or DNA or RNA sequence, that the indicated molecule is present in the substantial absence of other biological macromolecules, such as other proteins. The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above. "Isolated" and "purified" do not encompass either natural materials in their native state or natural materials that have been separated into components (e.g., in an acrylamide gel) but not obtained either as pure (e.g. lacking contaminating proteins, or chromatography reagents such as denaturing agents and polymers, e.g. acrylamide or agarose) substances or solutions.

Yet another aspect of the present invention concerns preparations of the subject antibodies, particularly pharmaceutical preparations. The antibodies of the present invention, or pharmaceutically acceptable salts thereof, may be conveniently formulated for administration with a biologically acceptable medium, such as water, buffered saline, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol and the like) or suitable mixtures thereof. The optimum concentration of the active ingredient(s) in the chosen medium can be determined empirically, according to procedures well known to medicinal chemists, and may depend on such as factors as intended route of administration, age and body weight of patient. As used herein, "biologically acceptable medium" includes any and all solvents, dispersion media, and the like which may be appropriate for the desired route of administration of the pharmaceutical preparation. The use of such media for pharmaceutically active substances is known in the art. Except insofar as any conventional media or agent is incompatible with the activity of the antibody, e.g., its specificity and/or affinity, its use in the pharmaceutical preparation of the invention is contemplated. Suitable vehicles and their formulation inclusive of other proteins are described, for example, in the book *Remington's Pharmaceutical Sciences* (Remington's Pharmaceutical Sciences. Mack Publishing Company, Easton, Pa., USA 1985). These vehicles include injectable "deposit formulations". Based on the above, such pharmaceutical formulations include, although not exclusively, solutions or freeze-dried powders of the antibody in association with one or more pharmaceutically acceptable vehicles or diluents, and contained in buffered media at a suitable pH and isosmotic with physiological fluids. For illustrative purposes only and without being limited by the same, possible compositions or formulations which may be prepared in the form of solutions for the treatment of proliferative disorders with an anti-cancer cell antibody of the present invention are given in U.S. Patent No. 5,218,094. In the

case of freeze-dried preparations, supporting excipients such as, but not exclusively, mannitol or glycine may be used and appropriate buffered solutions of the desired volume will be provided so as to obtain adequate isotonic buffered solutions of the desired pH. Similar solutions may also be used for the pharmaceutical compositions of the antibodies in isotonic solutions of the desired volume and include, but not exclusively, the use of buffered saline solutions with phosphate or citrate at suitable concentrations so as to obtain at all times isotonic pharmaceutical preparations of the desired pH, (for example, neutral pH).

Still another aspect of the present invention concerns assay kits that can be used for detecting an immunorecessive epitope(s) in a sample, for example. The assay kits generally provide an antibody for the immunorecessive epitope, derivatized with a label group that can be ultimately detected, as for example, by spectrophotometric techniques (including FACS) or radiographic techniques. For instance, the label can be any one of a number of radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors. To illustrate, the label group can be a functional group selected from the group consisting of horseradish peroxidase, alkaline phosphatase, ß-galactosidase, luciferase, urease, fluorescein and analogs thereof, rhodamine and analogs thereof, allophycocyanin, R-phycoerythrin, erythrosin, europiam, luminol, luciferin, coumarin analogs, ^{125}I, ^{131}I, ^{3}H, ^{35}S, ^{14}C and ^{32}P.

Assay kits provided according to the invention may include a selection of several different types of the subject antibodies. The antibodies may be in solution or in lyophilized form. In some embodiments, the antibodies may come pre-attached to a solid support, or they may be applied to the surface of the solid support when the kit is used. The labeling means may come pre-associated with the antibody, or may require combination with one or more components, e.g., buffers, antibody-enzyme conjugates, enzyme substrates, or the like, prior to use. Many types of detectable labels are available and could make up one or more components of a kit. Various detectable labels are known in the art, and it is generally recognized that a suitable label group is one which emits a detectable signal. Various label groups can be used, depending on the type of immunoassay conducted. Useful labels include those which are fluorescent, radioactive, phosphorescent, chemiluminescent, bioluminescent, and free radical. Also, the label groups may include polypeptides (e.g., enzymes or proteins), polymers, polysaccharides, receptors, cofactors, and enzyme inhibitors. Kits of the invention may also include additional reagent. The additional reagent can include blocking reagents for reducing nonspecific binding to the solid phase surface, washing reagents, enzyme substrates, and the like. The solid phase surface may be in the form of microtiter plates, microspheres, or the like, composed of polyvinyl chloride, polystyrene, or the like materials suitable for immobilizing proteins.

VI. Exemplary Applications of the Subject Method, and Antibodies Derived Therewith

The subject method of the present invention can be applied advantageously to the production of antibodies useful in purification, diagnostic, and therapeutic applications. In contrast to even the antibody display libraries which can be derived from immunized animals, the antibody libraries which can be generated by the subject method provide a greater population of high affinity antibodies to the immunorecessive epitope of interest, as well as establish a broader pool of display packages comprising antibodies specific for the immunorecessive epitope. With respect to the immunorecessive epitope, the more effective access of the antibody repertoire provided by the display libraries of the present invention allows more efficient enrichment to occur by, for example, affinity selection means.

As described above, immunorecessive epitopes can be defined in terms of the toleragen and immunogen used in the subtractive immunization step, and are therefore unique to the immunogen with respect to the toleragen. Thus, where the desired antibody is to distinguish between various cells of common or similar origin or phenotype, the cell to be specifically bound by an antibody of the present invention is used as an immunogen, while the related cells from which it is to be distinguished are employed as the toleragen. Table 1 provides exemplary systems of immunogen/toleragen sets which can be employed in the subject method to isolate antibodies which specifically bindepitopes unique to the immunogen. The choice of toleragen and immunogen can provide antibodies specific to, for example, tumor cell markers, fetal cell markers, and stem cell markers.

Likewise, the subject method can be used to generate antibodies which can discriminate between a variant form of a protein and other related forms of the protein by employing an immunogen comprising a variant protein, such as a mutant form of a protein or a particular isoform of a family of proteins, and a toleragen comprising the wild-type protein or alternate isoforms of the variant protein. The difference in determinants (i.e. the immmunorecessive epitopes) between the variant protein and wild-type (or other isoforms) will typically consist of only a few differences in amino acid residues (i.e. less than 15%, but preferably on the order of only one to three residues difference). For example, such combinations of immunogens and toleragens can be used in the present invention to derive antibodies which can specifically bind variant forms of oncoproteins or tumor suppressor proteins, as well as of hemoglobin, apolipoprotein E, LDL receptor, cardiac β-myosin, sodium or other ion channels, collagen, glucokinase, or transthyretin.

Table 1

Target Antigen	Toleragen	Immunogen
Fetal nucleated red blood cells	maternal erythroid cells	fetal erythroid cells
colon cancer	normal colon cells (e.g. epitheleal)	colon carcinoma cells
ApoE4	ApoE	ApoE4
Stem or embryonic nerve cell	differentiated nerve cell	embryonic nerve cell
hematopoietic cell	committed stem cell	hematopoietic stem cell
metastatic tumor marker	non-mestatic transformed cell	metstatic transformed cell
p53 mutant	wild-type p53	mutant p53

In an exemplary embodiment of the present invention, the subject method is employed to generate antibodies for a cell-type specific marker. As Example 2 illustrates, the present method can be employed to produce antibodies directed specifically to fetal cell-specific markers. For example, specific antibodies for markers of fetal nucleated red blood cells can be generated by the subject method employing maternal erythroid cells as a toleragen and fetal erythroid cells as an immunogen. When generated to distinguish between fetal cells and maternal cells, as by the specific recognition of epitopes on cell surface antigens such as of haematopoietic precursor cells, antibodies generated by the subject method can be used to separate fetal cells from maternal blood by, for instance, fluorescence-activated cell sorting (FACS). The isolated fetal cells, such as fetal nucleated erythrocytes, represent a non-invasive source of fetal DNA for prenatal genetic screening and offer a powerful and safe alternative to more invasive procedures than, for example, amniocentesis or chronic villus sampling.

In another embodiment, the present invention contemplates the generation of antibodies specific for a tumor cell-specific marker. As described in Example 1, the subject method can be employed advantageously to generate antibodies which are able to differentiate between normal cells and their transformed counterparts. Such antibodies may be suitable for both diagnostic and therapeutic uses. For example, antibodies can be selected in the present assay which detect cell-specific markers found on neoplastic or hyperplastic cells. Antibodies so obtained can be used to identify transformed cells and thereby used to diagnose cancers and tumors such as adenocarcinomas, papillomas, squamous and transitional cell carcinomas, anaplastic carcinomas, carcinoid tumors, mesotheliomas, hepatomas, melanomas, and germ cell tumors. These antibodies mays also be used to selectively destroy transformed cells, both *in vivo* and *in vitro*, such as through the

discriminatory activation of complement at the cell surface of a transformed cell bound by the antibody, or by delivery of toxins, or by delivery of nucleic acid constructs for gene therapy. For example, antibodies specific for colon cancer markers can be generated in the present invention by suing normal colon cells as a toleragen and cells derived from a colon carcinoma as an immunogen. In similar fashion, the subject method can be engaged to produce antibodies that specifically inhibit metastasis of highly metastatic tumor cells. Such antibodies, designed to recognize unique epitopes on highly metastatic variants of tumor cells (i.e. whose expression is elevated relative to non-metastatic variants), can be used to interfere with the function of cell surface proteins containing these epitopes in the metastatic cascade.

In similar fashion, where specific antibodies for stem or embryonic nerve cell markers are desired, the immunotolerance-derived antibody repertoires used in the subject method can be generated using a differentiated nerve cell as a toleragen and an embryonic nerve cell, such as a neural crest cell or uncommitted progenitor cell, as an immunogen.

For hematopoietic cell specific antibodies, the immunogen can comprise a hematopoietic stem cell, and the toleragen can be a committed stem cell.

In yet a further embodiment, the subject method can be applied to the generation of antibodies which can discern between variant proteins. Such antibodies can be used to distinguish various naturally occurring isoforms of a protein, as well as to detect mutations which may have arisen in a protein. In an illustrative embodiment, antibodies can be produced by the present invention which can be used in immunochemical assays for detecting cell transformations arising due to mutation of an oncogene or anti-oncogene. For instance, the subject method can be used to generate antibodies which discriminate between wild-type *ras* and a mutant form of *ras*. For example, useful antibodies for detecting *ras*-induced transformation of a cell can be generated by the subject method using a Ser-17→Asn variant of *ras* as an immunogen, and wild-type *ras* as a toleragen..

Likewise, diagnostically useful antibodies can be produced by the present invention which specifically bind and discriminate between wild-type and variant tumor suppressor proteins. For example, inactivating mutations of either the p53 or Rb tumor suppressors can lead to escape from cell senescence and lead to transformation. The subject method can be used to generate antibodies specific for a variant p53, the ability to distinguish between the wild-type and mutant forms arising through recognition of a unique epitope created by mutation, such as Arg-273→Cys, Tyr-163→Asn, Val-157→Phe, or Cys-238→Phe. Appropriate immunogen/toleragen sets would therefore include p53 mutants and wild-type p53.

The subject method can also be used to produce antibodies for detecting variant hemoglobin molecules, and which subsequently can be employed as diagnostic tools for

detecting hemoglobinopathies, such as sickle cell anemia and β-thalassemia. A large number of such abnormalities, most resulting from single-point mutations, have been observed as abnormal hemoglobins of embryonic, fetal, neonatal, and adult disorders (see, for review, Huisman (1993) *Baillieres Clin Haematol* 6:1-30). Therefore, antibodies to unique epitopes of hemoglobin variants can be of great use in detecting and quantitating both normal and abnormal hemoglobin levels.

Where the immunogen is apolipoprotien E4 (ApoE4) and the toleragen comprises other ApoE isoforms, specific antibodies can be isolated by the subject method which can be used to measure ApoE4 levels in plasma or serum of a patient. The presence of the ApoE4 variant has been linked to increased susceptibility to Alzheimer's disease (Strittmatter et al. (1993) *PNAS* 90:8098-8102) as well as significant impact on variation of cholesterol lipid and lipoprotein levels in individuals (Rall et al. (1992) *J. Intern. Med.* 231:653-659; and Weisgraber et al. (1990) *J. Lipid Res.* 31:1503-1511). In similar fashion, specific antibodies to other ApoE isoforms can be generated, including antibodies which can specifically bind ApoE2 or ApoE5.

Other exemplary embodiments include the generation of specific antibodies for LDL receptor variants which can be useful, for example, in predicting risk of diagnosing familial hypercholesterolemia; specific antibodies to cardiac β-myosin variants, which can be used to diagnose hypertrophic cardiomyopathy; specific antibodies to variant forms of sodium or ion channels, such as which arise in congenital hyperkalemic periodic paralysis; antibodies to collagen variants, such as Cys-579 collagen, which can be indicative of a predisposing factor in risk of familial osteoarthritis; specific antibodies to a variant of glucokinase, such as which arise in non-insulin-dependent diabetes mellitus; and antibodies specific for a mutant of transthyretin, such as which might arise in familial amyloidotic polyneuropathy.

VII. Antibodies Specifically Reactive with Fetal Cell/Cancer Cell Epitopes

As described in detail below, the subject method has been applied advantageously to the development of antibodies for cell-surface markers of fetal cells and transformed cells. In contrast to the use of conventional hybridoma methods, or even pre-immunized phage display libraries, practice of the subject method can yield a library of antibodies which are amenable to very rapid enrichment. This invention represents the first instance that antibodies specific for unknown/unisolated cell-surface antigens have been generated using a combinatorial display library. Indeed, initial experimentation using V-gene libraries derived from animals immunized with the immunorecessive epitope, but not tolerized to background epitopes (in contrast to the present method), suggests that the subject antibodies are attainable only with

great difficulty and expense, and perhaps not at all, by prior art combinatorial display techniques.

To illustrate, Figure 5A reveals the rapid enrichment of specific antibodies from the immunotolerized V-gene library. By comparison, Figure 5B demonstrates that phage libraries prepared by prior art techniques (non-tolerized #1 and #2) do not show significant enrichment from one round of panning to the next (compare tolerized to non-tolerized #1 and #2). Likewise, as set out in more detail below, despite several years of investigating hybridomas, even those generated using B-cells derived by immunotolerization protocols, antibodies that discriminate between fetal and maternal blood cells with only the same approximate performance as anti-CD71 antibodies were obtained.

The subject method, on the other hand, provides a library containing a rich source of high affinity antibodies which permit detection of specific antibodies by, for example, panning on live cells, FACS assays or cell based ELISA. To further illustrate, the appended examples describe that individual antibody display packages were enriched 5000 to 3,600,000 fold in only a single round of selection. DNA sequence analyses of particular isolates depict a remarkable history of affinity maturation of both heavy and light chains, suggesting an unexpectedly efficient access to the immunological repertoire.

Furthermore, in addition to hastening antibody maturation, and perhaps causing such enhanced maturation, the instant method enables selection of antibodies having both discriminating specificity and high binding affinity for an immunorecessive epitope. Indeed, comparison of antibodies isolated by the subject method with antibodies available through the use of prior art techniques reveals that the combinatorially-derived antibodies of the present invention tend to be orders of magnitude better with respect to each of specificity and affinity relative to antibodies available in the prior art.

The genes for three of the antibodies which demonstrate both desirable specificity and binding affinity have been sequenced. As described in Example 2, the F4-7 and H3-3 antibodies were originally isolated with a panning regimen including fetal liver cells. Further characterization of the H3-3 antibody confirmed that this antibody recognized fetal blood cells of early gestational age (e.g., <16 weeks), but also stained fetal cells of later gestational ages, albeit less well. This probably reflects the use of fetal liver, which consists predominantly of the earliest blood cell precursors, for both immunization and enrichment. However, it is demonstrated below that the population of antibodies enriched from the library could be biased to select antibodies specific for epitopes present on fetal blood cells of later gestational ages. One of the isolates, FB3-2, was characterized and found to have an extraordinarily low background staining level on adult blood cells (e.g., less than 0.1%). A guide to the nucleic acid and amino acid sequences for each of these clones is provided in

Table 2, and the overall structure of the variable region for each of the heavy and light chains are provided in Figures 8A and 8B.

Table 2
Nucleotide and Amino Acid Sequences for Anti-Fetal Antibodies

Antibody	H.C. nucleotide	H.C. amino acid	L.C. nucleotide	L.C. amino acid
FB3-2	SEQ ID No. 50	SEQ ID No. 51	SEQ ID No. 52	SEQ ID No. 53
F4-7	SEQ ID No. 54	SEQ ID No. 55	SEQ ID No. 56	SEQ ID No. 57
H3-3	SEQ ID No. 58	SEQ ID No. 59	SEQ ID No. 60	SEQ ID No. 61

The antibodies isolated by the present method, derived from a V-gene library of an immunotolerized animal, are not apparently available by other prior art techniques and in fact displayed performance characteristics which greatly surpassed those obtained by previous methods. In contrast to the antibodies achieved by the subject method, employing an identical immunotolerization step, but coupled instead with the use of hybridoma techniques, only a few antibodies which showed fetal cell selectivity were obtained. The specificity for one of the best of these antibodies, "anti-Em", is shown in Table 3. Fetal cell selective antibodies isolated by other groups using other hybridoma technologies were also compared. As is understood in the art, anti-CD71 antibodies are believed to be among the best of the fetal cell specific antibodies. However, as demonstrated in Table 3 (see also example 4), antibodies generated by the instant method perform with superior qualities relative to each of the antibodies obtained by immunotolerance (anti-Em) and hybridoma (anti-CD71) techniques.

Table 3
Comparison of hybridoma-derived antibodies with immunotolerized/phab-derived antibodies

Antibody	Amt Used	Cell Type Stained	% Positive	Specificity
Anti-Em	5.0 μg	fetal liver	50.0%	2.5 fold
	5.0 μg	maternal PBMC	20.0%	
H3-3	0.25 μg	fetal liver	79.5%	>125 fold
$F(ab')_2$	0.25 μg	maternal PBMC	below background	
FB3-2	0.25 μg	fetal liver	85.4%	125 fold
$F(ab')_2$	0.025 μg	fetal liver	80.0%	
	0.25 μg	maternal PBMC	0.68%	
anti-CD71 (Beckton-Dickinson)	1.0 μg	fetal liver	83.1%	7.7 fold
	1.0 μg	maternal PBMC	10.8%	

Each of the anti-Em and anti-CD71 antibodies are considered to be of excellent specificity with respect to anti-fetal cell antibodies derived by methods in the prior art. Yet, as Table 3 illustrates, the level background binding to maternal peripheral blood mononuclear cells (PBMC) is many times higher for these antibodies relative to the background staining of maternal cells using the subject antibodies. Consequently, although the anti-Em, anti-CD71 antibodies and the like stain fetal cells very well, their background staining on maternal blood of greater than 5 percent provides substantial room for improvement of antibodies useful for retrieving a very small population of fetal blood cells from maternal blood samples.

One estimate of fetal cell concentrations in maternal blood provides 1 fetal cell in 100,000 to 1 in one million (e.g., %0.001 to %0.0001) adult nucleated blood cells. The performance attributes on the antibodies derived by the subject method suggests that these antibodies are specific enough for use in purifying fetal cells from maternal samples. To further demonstrate the improved performance of antibodies isolated by the subject method, relative to antibodies known in the art, 400 male fetal cells were spiked into 1 million maternal cells, and the mixture stained with fluorescein-conjugated H3-3 Fab. FACS sorting to recover stained cells, followed by *in situ* hybridization to detect Y chromosomal DNA, demonstrated a recovery of 75% of the male cells at almost 40% purity (300 male cells of 800 total cells recovered). The best results obtained for either of the anti-Em or anti-CD71 antibodies described above approach the same percentage of recovery of fetal cells, but at orders of magnitude lower purity (e.g., a few hundred fetal cells amongst a background of 100,000 maternal PBMC).

Another feature of the antibodies derived from the subject method, which feature also apparently exceeds the antibodies of the prior art, pertains to the binding affinity of these antibodies for fetal cell-bound antigens. As described in Example 3, the affinity of the H3-3 and FB3-2 antibodies was determined against human erythro-leukemic (HEL) cells ("HEL scatchard assay"). In each instance, the association constant (K_a) exceeded 10^9. For instance, monomeric H3-3 and FB3-2 Fab' fragments displayed association constants of $6x10^{10}M^{-1}$ and $8x10^{10}M^{-1}$ respectively. Dimeric forms of the recombinant antibodies had even greater binding affinities, with K_as of $5x10^{12}M^{-1}$ for H3-3 and $1x10^{12}M^{-1}$ for FB3-2 respectively.

As a result of the inventors' discovery, it is now possible to provide a reproducible and predictable method for isolating antibodies immunoreactive with immunorecessive epitopes, which antibodies are characterized by specificity and/or affinity for a corresponding antigen which exceed those presently attainable by either hybridoma or by phage display technologies. Accordingly, in one aspect of the invention, the subject method makes

available antibodies specific for immunorecessive epitopes, in which antibodies are characterized by association constants for the immunorecessive epitopes which are greater than 10^6M^{-1}, preferably greater than 10^8M^{-1}, more preferably greater than about 10^9M^{-1}, and even more preferably greater than $10^{10}M^{-1}$, $10^{11}M^{-1}$, or $10^{12}M^{-1}$, e.g., K_a in the range of $10^{10}M^{-1}$ to $10^{13}M^{-1}$.

In another aspect of the invention, the subject method accommodates the isolation of antibodies which have a low level of background staining. The relative specificity of these antibodies can be several fold, if not orders of magnitude, better than combinatorial and hybridoma generated antibodies, particularly with respect to antibodies for cell surface epitopes. For instance, the subject method can provide antibodies which have no substantial background binding to other related cells, e.g., relative specificities greater than 10 fold binding to the target cells over background binding to the related cells. As demonstrated, antibodies can be generated which do not substantially cross-react with other epitopes, preferably having specificities greater than 20 fold over background, more preferably 50, 75 or 100 fold over background, and even more preferably more than 125 fold over background. For example, anti-fetal cell antibodies generated by the instant method, exemplified by the FB3-2 and H3-3 antibodies, were tested by fluorescence-activated cell sorting ("FACS efficiency assay") and were each demonstrated to have relative specificities greater than 125 fold over background. In contrast, the anti-CD71 and anti-Fe antibodies were found to have relative specificities of 7.7 and 2.5 fold over background, respectively. Furthermore, the specificity of fetal cell specific antibodies produced by the subject method can also be characterized in terms of a background staining of maternal cells relative to antibodies of the prior art, such as anti-CD71 antibodies. For instance, the subject antibodies preferably stain two times less non-fetal cells relative to an anti-CD71 antibody, more preferably at least five times less, and even more preferably at least twenty times less than an anti-CD71 antibody. Such comparisons can be made using standard immunoassays, such as the FACS efficiency assay of Example 4. Exemplary anti-CD71 (e.g., anti-Transferrin receptor) antibodies include the 5E9 antibody (ATCC HB21), the L5.1 antibody (ATCC HB84) and the L01.1 antibody (Beckton Dickinson Catalog No. 347510).

With respect to the specific antibodies which have been sequenced, namely H3-3, F4-7 and FB3-2, it is emphasized that, as described above, each antibody can be further engineered without departing from the purpose and intent of the present invention. Accordingly, a chimeric FB3-2 antibody can be generated which includes the variable regions from the heavy chain (residues E1-S121, SEQ ID No. 51) and light chain (residues D1-K111, SEQ ID No. 53). Likewise, chimeric F4-7 antibodies can be provided which include the heavy chain (residues E1-S121, SEQ ID No. 55) and light chain (residues D1-K111, SEQ ID No. 57) variable regions from the F4-7 antibody described below. Furthermore, chimeric H3-

3 antibodies are also contemplated, as for example antibodies including the variable regions from the heavy chain (residues E1-S115, SEQ ID No. 59) and light chain (residues D1-K111, SEQ ID No. 61) of the H3-3 antibody.

In similar fashion, chimeric antibodies can be generated including heavy and light chain variable regions, each represented by the general formula: FR(1)-CDR(1)-FR(2)-CDR(2)-FR(3)-CDR(3)-FR(4), wherein CDR(1), CDR(2) and CDR(3) represent complementarity determining regions from the subject antibody, and FR(1), FR(2), FR(3) and FR(4) are framework regions from a second antibody. For example, chimeric FB3-2 antibodies can be generated which include a heavy chain in which CDR(1) is SYWLE, CDR(2) is EILFGSGSAHYNEKFKG and CDR(3) is GDYGNYGDYFDY, and a light chain in which CDR(1) is RASQSVSTSRYSYMH, CDR(2) is FASNLES and CDR(3) is HSWEIPYT. Likewise, a chimeric F4-7 antibody can be made including a heavy chain in which CDR(1) is SSWLE, CDR(2) is EILFGSGSAHYNEKFKG and CDR(3) is GDYGNYGDYFDY, and a light chain in which CDR(1) is RVRQSVSTSSHSYMH, CDR(2) is YASNLES and CDR(3) is HSWEIPYT. In similar fashion, chimeric H3-3 antibodies can be provided, which antibodies include a heavy chain having a CDR(1) of DYYMY, a CDR(2) of TISDDGTYTYYADSVKG and a CDR(3) of DPLYGS, and a light chain in which CDR(1) is RSSQSLVHSNGNTYLH, CDR(2) is KVSNRFS and CDR(3) is SQSTHVLT. In each instance, the associated framework regions (FR1-FR4) can be derived from an unrelated antibody, preferably a human antibody.

The present invention further pertains to methods of producing the subject recombinant antibodies. For example, a host cell transfected with nucleic acid vectors directing expression of nucleotide sequences encoding an antibody (or fragment) can be cultured under appropriate conditions to allow expression of the antibody to occur, and if required, assembly of a heavy/light chain dimer. The antibody may be secreted and isolated from a mixture of cells and medium containing the recombinant antibody. A cell culture includes host cells, media and other by-products. Suitable media for cell culture are well known in the art. The recombinant antibody peptide can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying antibodies, including protein-A:sepharose and ion-exchange chromatography, gel filtration chromatography, ultrafiltration and electrophoresis. In a preferred embodiment, the recombinant antibody is a fusion protein containing a domain which facilitates its purification, such as a GST fusion protein or a poly(His) fusion protein.

This invention also pertains to a host cell transfected to express a recombinant form of the subject antibody. The host cell may be any prokaryotic or eukaryotic cell, and the choice can be based at least in part on the desirability of such post-translation modifications as glycosylation. Thus, a nucleotide sequence derived from the cloning of an anti-fetal cell or

anti-oncogenic cell antibody by the subject method, encoding all or a selected portion of the variable region, can be used to produce a recombinant form of an antibody via microbial or eukaryotic cellular processes. Ligating the recombinant antibody gene into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures used in producing other well-known proteins, e.g., insulin, interferons, human growth hormone, IL-1, IL-2, as well as other recombinant antibodies. Similar procedures, or modifications thereof, can be employed to prepare the subject antibodies by microbial means or tissue-culture technology in accord with the subject invention.

Preferably, the cell line which is transformed to produce the recombinant antibody is an immortalised mammalian cell line, which is advantageously of lymphoid origin, such as a myeloma, hybridoma, trioma or quadroma cell line. The cell line may also include a normal lymphoid cell, such as a B-cell, which has been immortalised by transformation with a virus, such as the Epstein-Barr virus. Most preferably, the immortalised cell line is a myeloma cell line or a derivative thereof.

The recombinant antibody gene can be produced by ligating nucleic acid encoding the subject antibody protein, or the heavy and light chains thereof, into vectors suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vectors for production of recombinant forms of the subject antibody include plasmids and other vectors. For instance, suitable vectors for the expression of an antibody include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach *et al.* (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, an antibody is produced recombinantly utilizing an expression vector generated by sub-cloning the coding sequences of the variable regions for each of the heavy and light chain genes of the H3-3 or FB3-2 antibodies.

The preferred mammalian expression vectors contain both prokaryotic sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived

vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papillomavirus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant antibody by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the ß-gal containing pBlueBac III).

VIII. Exemplification

The invention now being generally described, it will be more readily understood by reference to the following examples which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention.

As described below, the subject method has been applied advantageously to the development of antibodies for cell-surface markers of fetal and transformed cells. In contrast to the use of conventional hybridoma methods, or even pre-immunized phage display libraries, the present invention can yield a remarkable library of antibodies which are amenable to very rapid enrichment. In an exemplary embodiment described in the Examples below, individual antibody display packages were enriched 5000 to 3,600,000 fold in only a single round of selection. DNA sequence analyses of particular isolates gave a remarkable history of affinity maturation of both heavy and light chains, suggesting an unexpectedly efficient access to the immunological repertoire.

I. Materials and Methods

Except where indicated otherwise, recombinant DNA methods and microbiological techniques were carried out as described by Sambrook, J. et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

The materials and methods described below were used to generate the antibody display libraries described in both Example 1 and Example 2.

Biochemicals

DNA modifying enzymes were obtained from New England Biolabs (Beverly, MA) and used under conditions recommended by the suppliers. Taq polymerase was obtained from Perkin Elmer (Norwalk, CT). A set of DNA fragments (1 Kb ladder) obtained from Life Technologies (Gaithersburg, MD) was used as a standard for molecular weight of DNA fragments by agarose gel electrophoresis. DNA primers were custom synthesized by Genosys, Inc. (The Woodlands, TX) or Cruachem, Inc. (Sterling, VA). Deoxyadenosine 5'[α -(^{35}S)thio]triphosphate was purchased from New England Nuclear (Boston, MA). Polyclonal biotinylated anti-M13 antibody was obtained from 5 prime-3 prime (Boulder, CO). Streptavidin-Alkaline phosphatase and Polyclonal goat anti-mouse kappa-alkaline phosphatase were from Fisher Biotech (Pittsburgh, PA).

Bacterial strains and culture

E. coli strains XL-1, SolR, and LE392 were obtained from Stratagene (LaJolla, CA). Lambda phage resistant XL-1 was isolated by standard methods and is described in this work. *E. coli* was grown to stationary phase at 30 or 37°C with shaking in Erlenmeyer flasks filled to one-tenth their nominal capacity with LB, SOB, 2X YT, NZY medium (Sambrook, 1989) or TB medium:0.1 M KH2PO4 buffer, buffer, pH 7.5 containing 12 g bacto-tryptone, 24 g yeast extract, and 5.04 g glycerol per liter (phosphate buffer was autoclaved separately). For growth of bacteria on solid media, agar (Difco, Detroit MI) was added to a final concentration of 2% (wt/vol.). Glucose supplement was to 0.5% Carbenicillin, chloramphenicol, and kanamycin were added when necessary to 50, 30, and 50 ug/ml, respectively.

Vectors and bacteriophage

The *E. coli* cloning vector, lambda SurfZap™ and helper phages ExAssist™ and VCS M13 were obtained from Stratagene.

Cell Preparation

Whole blood from non-pregnant individuals was obtained from Interstate Blood Products (Tennessee). Adult peripheral blood mononuclear cells ("PBMC") were prepared by standard Ficoll-Hypaque gradient techniques. Fetal blood mononuclear cells were prepared from fetal liver obtained from abortuses at 12-20 weeks gestation, at which age the liver is the principal hematopoietic organ. Cells were freed from surrounding connective tissue by passage through sterile microscreens in the presence of sterile Ca-Mg-free PBS. The resulting cell suspension was diluted up to 20 ml in PBS and the blood mononuclear cell

fraction obtained by standard Ficoll-Hypaque gradient centrifugation. After recovery from the Ficoll interface, both adult and fetal cells were washed twice in sterile Ca-Mg-free PBS the resuspended in the PBS at $2x10^7$ cells per ml.

Tolerance Immunizations

The use of cyclophosphamide tolerance with intact, fixed, cells is a well known technique in the art. The present procedure employed unfixed cells immediately after isolation from whole blood, bone marrow or fetal liver, avoiding the issue of alteration of antigens by chemical and/or physical processing of the cells. Cyclophosphamide was obtained from Sigma chemical and reconstituted at 10 mg/ml sterile saline. The tolerization procedure used was essentially that of Matthew et al. (1987) *J Immunol Methods* 100:73-82.

An alternating schedule of tolerization and immunizations was set up as follows: female Balb/c mice at 6 weeks of age were injected intra-peritoneally ("I.P.") with $1x10^7$ adult PBMC in 500 ul PBS. The adult PBMC injection was followed 10 minutes later by I.P. injection of cyclophosphamide at 100 mg/kg. The cyclophosphamide was repeated at 24 and 48 hours. After an additional 14 days, the tolerization was repeated.

Three weeks later, the mice were immunized with fetal mononuclear blood cells by I.P. injection of $1x10^7$fetal cells in 500 ul PBS. After an additional 2 weeks, the mice were once again tolerized with adult PBMC as described for the first round of tolerization. Finally, three weeks later, the mice were again immunized with fetal blood mononuclear cells by I.P. injection of $1x10^7$ fetal cells in 500 ul PBS. The fetal cell immunization was repeated in 24 and 48 hours. After an additional 24 hours, the mice were sacrificed. The spleens were harvested and immediately frozen in liquid nitrogen.

Isolation of RNA and cDNA synthesis

Total RNA was isolated from spleens or from Hybridoma cell lines using standard methods (Chomczynski, U.S. Patent No. 4,843,155). RNA preparations were stored in RNAase free water (Sambrook, J. *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)) at -70°C until use. A Superscript pre amplification kit from Life Technologies was used to prepare first strand cDNA as recommended by the supplier.

Isolation of DNA

Isolation of plasmid DNA from *E. coli* for DNA sequence or restriction analyses was by alkaline lysis (Birnboim and Doly, 1979). Bulk preparation of plasmid DNA was carrier out using nucleobond column chromatography as described by the manufacturer Macherey-Nagel (Duren, Germany). All cultures for isolation of plasmid DNA from *E. coli* clones

containing antibody clones were grown overnight with shaking at 37°C in 2xYT medium containing 0.5% glucose and 50 ug/ml carbenicillin.

PCR amplification of antibody kappa chain and IgG1 heavy chain coding regions

A set of degenerate primers, shown in Figures 1A and 1B, was designed to minimize bias toward limited sets of PCR products from the repertoire of antibody coding regions encoded in spleenic mRNA, as well as to amplify >90% of the mouse kappa chain and heavy chain Fab encoding sequences. Amplifications of kappa chain or heavy chain coding sequences were accomplished using 5 separate primer pairs for each. The primers also contained restriction enzyme site to allow the ligation of the light and heavy chain PCR products into a bacterial Fab expression cassette suitable for insertion in the Surf-Zap vector (Stratagene). PCR reactions were carried out in an Automated BioSystems temp-cycler (Essex, MA) using the following protocol. Generally, 5-10 ug of total spleenic or hybridoma RNA was converted to cDNA using a Superscript first strand synthesis kit (BRL). 0.5-1 ug of first strand cDNA in 100 ul of buffer containing 10 mM Tris-HCl, pH8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin with the appropriate primer pair (see Figures 1A and 1B) was incubated for 5 min. 98°C, cooled to 60°C, and 2 U of Taq polymerase was added. Products were then amplified 35 cycles with the following four temperature profile: 72°C for 120 seconds, 90 seconds at 54°C, and 30 seconds at 45°C. After 35 cycles the samples were incubated an additional 10 min. at 72°C. to ensure complete product polymerization. It is important that individual reactions be adjusted to yield approximately 1-2 ug of 0.7 Kb PCR product, with the minimal number of cycles (usually 30-38 cycles).

Assembly of Fab expression cassettes from kappa and heavy chain PCR products

1-2 ug of each PCR product from five separate reactions were combined to generate a kappa chain and separate heavy chain product pool. The pools were then purified by first removing protein and debris with a PVDF spin filter (Millipore) followed by removal of low molecular weight components using a 30,000 MW cut off spin filter as recommended by the supplier (Millipore). Approximately 5 ug of each product pool was digested in 300 ul of SfiI buffer with 50 units of SfiI for 2 h at 50°C. Enzyme and small end fragments generated by SfiI digestion were removed with the spin column procedure described above. SfiI digested light chain products were ligated to SfiI digested heavy chain products (approximately 2 ug each) in a 50 ul volume overnight at 4°C. The ligation mixture was then purified with spin columns as above and digested with 50 units each of NotI and SpeI restriction enzymes in 100 ul. The digestion products were resolved by agarose gel electrophoreses and the 1.4 kb kappa chain heavy chain encoding dimer was purified using Gene Clean II (Promega) as recommended by the supplier.

A simpler more reliable method for construction of the Fab' expression cassette is diagrammed in Figure 2. Approximately 10 ng of kappa and heavy chain product pools from above were amplified with primers designated in Figure 2 (and shown in Figures 1A and 1B) to give 3' kappa and 5' heavy chain sequence which when treated with T4 polymerase in the presence of dTTP yielded an 8 base compatible overhang allowing highly efficient oriented ligation of kappa and heavy chain sequences. PCR products were purified as described above. Approximately 2 ug of each product was treated separately at 37°C for 1 h in a 50 ul volume containing 5 units T4 polymerase, 5 mM dTTP. Products were purified as for PCR products, and approximately 500 ng of each product was ligated at room temperature for 3 h in a 25 ul volume of ligation buffer (Promega) containing 2 units of DNA ligase.

Fab encoding dimer from either method were amplified under standard conditions using a 5' kappa chain primer and 3' heavy chain primer shown in Figure 2 except that annealing was at 55°C for 1 min., and the extension time was extended to 4 min. at 72°C. Generally 12-25 cycles under these conditions yielded approximately 1-2 ug of 1.4 kb kappa-heavy chain dimer. This product was purified using spin columns as described above and then digested in a200 ul volume containing 75 units each of Not I and Spe I restriction enzymes. Digestion products were purified as described above except that a 100,000 MW spin column (Amicon) was used to more efficiently remove primers from the digestion products. Purified 1.4 kb dimers were stored at 4°C until use.

Construction of variegated Fab clone banks

Ligation of Not I-Spe I digested 1.4 kb Fab encoding fragments was as follows: 0.2 ug of digested products was ligated to 2 ug of lambda surf-zap arms in 10 ul of Promega ligation buffer containing 3 units of T4 ligase overnight at 4°C. Aliquots of the ligation mixture were then packaged into lambda heads using a Giga-pack Gold packaging kit as recommended by the supplier (Stratagene). Packaging reactions were titered on *E. coli* LE392 and pooled to yield a primary library. This primary library was then amplified in *E. coli* LE392 using conventional methods. Generally $5x10^9$ *E. coli* XL1 cells were infected in 10 ml of 10 mM $MgSO_4$ with 107 invitro packaged SURF-ZAP primary clones for 10 min. at 37°C. The infected cells were added to 100 ml of NZY top agarose at 50°C. The mixture was immediately plated onto two 20x20 cm plates containing NZY agar, allowed to solidify, and then incubated for 8-16 h. The amplified library was harvested by rocking with an overlay of 25 ml of SM buffer of 2 h.

Generation of Phage antibody clone banks

A Phagemid clone bank was rescued from the primary lambda SURF-ZAP library by super infection with M13 exassist helper phage essentially as recommended by Stratagen. Generally 10^{11} *E. coli* XL1 cells were infected with 10^{10} lambda clones from our amplified

surf-zap library and 10^{12} Exassist M13 phage. After growth for 3.5 h in LB or TB medium, the cells were removed by centrifugation. The exassist rescued library was treated for 70°C for 20 min. and then stored at 4°C.

Phage antibodies were generated by infection of *E. coli* SOLR 1:1 with rescued phagemid to generate a population of carbenicillin resistant antibody clone containing cells representing a 10-100 fold excess over the primary library size. Transduced cells were grown to early log phase in TB medium containing carbenicillin, and then infected with a ten fold excess of VCS M13 helper phage to cells. After 1 h at 37°C, kanamycin was added and the culture was incubated at 30°C with shaking until early stationary phase. Cells were removed by centrifugation, and phage antibodies were recovered from the supernatant by harvested by centrifugation, dissolved in 1 ml of TE buffer and then PEG precipitated a second time. Phage antibodies were dissolved in 1 ml TE or PBS buffer and stored at 4°C.

Enrichment of cell surface binding phabs on whole cells

Cell specific phage antibodies were isolated by enrichment on whole cells. Cells were prepared for enrichment by washing twice in blocking buffer (0.1% hydrolyzed casein, 3% BSA, in Hanks Buffered Salt Solution). For the first round of enrichment 10^{11} phage antibodies in 200 ul of blocking buffer were added to 10^6 cells and incubated on ice for 1 h. Non-specific phage antibodies were then removed by washing 8 times with cold blocking buffer. Cells were harvested after each wash by centrifugation at 3500 rpm in an Eppendorf micro centrifuge. Cell surface bound phage antibodies were then eluted in 500 ul of 0.2 M Glycine pH2.2 containing 3 M urea and 0.5% BSA. Debris was removed by centrifugation, and the supernatant was neutralized by addition of 50 ul of 1 M Tris pH 9.5. Urea and buffer components were removed with three buffer changes using an amicon 100,000 MW cut off spin column. Enriched populations of phage antibodies were titered on XL1 cells, and then amplified by the following protocol. Eluted phage antibodies in 200 ul SM buffer were added to $5x10^9$ XL1 plating cells in 1 ml of 10 mM $MgSO_4$ and incubated for 10 min. at room temperature. Infected cells were then used to inoculate 100 ml of TB broth in a 2 L flask and incubated at 30°C with shaking. After 1 h of incubation kanamycin and carbenicillin were added to 50 ug/ml. Incubation was continued with shaking at 30°C until early stationary phase (Increase in O.D.600 remained the same at two consecutive time points). Cells were removed by centrifugation at 12000 rpm in a Dupont/Sorvall SS34 rotor for 10 min. Phage antibodies were then purified by PEG precipitation. The stringency of washes in subsequent rounds of enrichment were increased by reducing the amount of phage antibodies loaded to 10^{10}, increasing the number of washes (10-20 washes) and adding a 100 mM citrate buffer wash (pH 4.5 or pH3.5) before elution with Glycine-Urea.

Preparation of Individual phage antibodies for analyses of binding specificity

E. coli XL1 was infected with dilutions of phage antibody pools and plated on LB medium containing 0.5% glucose and 50 ug/ml carbenicllin. 20 mm culture tubes containing 2 ml of 2xYT medium with 50 ug/ml carbenicllin were inoculated with isolated colonies and grown overnight at 30°C with shaking. The following morning 1 ml of culture was gently shaken at 37°C for 1 h and then infected with 10^{11} M13 VCS phage. A 250 ml flask containing 25 ml of TB broth with carbencillin was inoculated with the VCS infected culture and shaken at 30°C for 1 h, kanamycin was added to 50 ug/ml and incubation was continued until early stationary phase (O.D. 600 nm = 5-12). Cells were removed by centrifugation, and phage antibodies were purified by PEG precipitation as previously described. The phage antibodies were dissolved in 200 ul TE (pH7.5) buffer.

Sequencing of Antibody Isolates

Individual isolates, such as the H3-3, FB3-2 and F4-7 clones, were isolated and each of the heavy and light chain inserts were sequenced using primers based on 3' and 5' flanking sequence of the Surf-Zap vector using standard protocols (see *Current Protocols in Molecular Biology*, eds. Ausubel et al. (John Wiley & Sons: 1992); and *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press:1989)).

Flow cytometric assay of phage antibody binding to whole cells

A flow cytometry protocol was devised for the testing of phage antibody binding to surface markers on whole cells. $1x10^6$ adult or fetal mononuclear cells were dispensed into a 2 ml microtube and washed with blocking buffer as in the phage enrichment procedure. For initial assay, $2x10^{10}$ phage were added to the washed cells and the volume brought to 100 ul with casein/BSA/HBSS. The phage were incubated with the cells for one hour at 4°C. The phage-cells were washed three times with 1 ml blocking buffer. Biotinylated sheep anti-phage polyclonal antibody (5 Prime -2 Prime) was added to the phage-cells at 5-7.5 ul per sample, optimized for each lot of polyclonal antibody. Volume was once again brought up to 100 ul with blocking buffer. The anti-phage was incubated 90 minutes at 4°C. Excess anti-phage was removed by washing three times with blocking buffer. Streptavidin-FITC (Jackson Immunoresearch) was diluted 1:50 in Ca-Mg-free PBS and 250 ul added to each sample of phage-cells. After a 30 minute 4°C incubation, the phage-cells were washed twice with blocking buffer and fixed by adding 400 ul 0.5% formaldehyde.

All samples were analyzed by flow cytometry. Fluorescence background was determined by using a non-display phage (VCS M13) as a negative control. Intensity of

FITC fluorescence above background on each cell was directly proportional to the number of specifically bound phage.

Relative binding activity of each clone was determined by evaluation of two parameters: (1) scatter pattern vs. intensity of fluorescence, for determination of relative cell surface epitope number and uniformity of expression for each phage clone, with higher, more uniform, numbers being most desirable; (2) titration of phage and retention of fluorescent binding intensity - for determination of relative phage antibody affinities.

In a variation of the above binding assay, soluble antibody ("Fab") was tested for activity. For the Fab fragments, the anti-phage/streptavidin-FITC was replaced by a goat anti-IgG-FITC $F(ab')_2$ polyclonal antibody (TAGO) that recognized the κ chain of the Fab fragments. 30 ul of the goat anti-IgG-FITC diluted 1:10 in 2.5% normal human serum was used per sample. The dilution in human serum ensured that any cross-reactivity of the polyclonal with human blood cell antigens would be minimized.

II. Example 1: Enrichment of phage antibodies on cancer cells.

A combinatorial phage display library of IgG1 and kappa chain derived Fabs containing $6x10^7$ primary clones was constructed from a mouse which had been tolerized with adult human blood and immunized with fetal liver cells. In order to minimize clonal bias due to individual growth characteristics, cultures containing antibody clones or pools of clones were always in rich media (TB or 2xYT containing 1% glucose). In addition, cultures used to produce phage antibodies were harvested as close to peak growth as possible since binding activity was found to fall beyond the start of stationary phase of growth.

The human erythro-leukemic cell line (HEL) carries onco/fetal cell surface markers also found on fetal liver cells. This characteristic and the ability to culture this cell made it a reliable source of cells to develop methods for enrichment of cell line specific antibodies from the above phage library. The binding of phage antibody pools enriched on this cell line (HEL) are shown in Figure 3.

These results demonstrated an increase in binding from 2.5×10^{-7} to 1.25×10^{-3} after 3 rounds of enrichment, reflecting a 4 log increase in phage binding to the cell surface. We tested the specificities of phage antibodies prepared from 16 independent isolates on HEL, Raji, and Adult blood cells by fluorescence activated cell sorting. Specific binding of phage antibody to the cell surface was detected by biotin-labeled anti-phage antibody followed by fluorescence conjugated streptavidin. Results from these assays shown in Figure 4 demonstrated that at least 6 out of the 16 isolates (indicated by astericks) from enrichment 3

(i.e. phage isolated after three interative pannings) appeared specific for determinants found only on HEL cells.

In order to evaluate how diverse the population of antibodies being enriching on whole cells was, the DNA sequence encoding the regions including CDR3 from these six phage antibody isolates was determined. Of the six HEL specific isolates there was only a single duplicate. This result demonstrated that the directed isolation of HEL cell specific antibodies by enrichment on whole cells had been achieved.

III. Example 2: Enrichment of phage antibodies on fetal cells.

To maximize the chances of isolating fetal cell specific clones, the phage antibody library described in Example 1 was pre-absorbed on adult nucleated blood prior to each enrichment cycle on fetal liver cells in addition to enrichments without pre-adsorption. The results of sequential rounds of pre-adsorption and enrichment on fetal liver cells are shown in Figure 5A. The increase in the percentage of phage antibodies binding to fetal liver cells indicated enrichment for fetal cell binding phage antibodies.

Remarkably, after only a single round of enrichment on fetal cells, it was observed that pure populations of fetal cell binding phage antibodies had been isolated. This was demonstrated by characterization of a sampling of random isolates, from each stage of enrichment, by DNA sequence analysis in combination with an assay for binding of individual isolates to fetal cells by FACS. The results of these experiments are shown in Figure 7. DNA sequencing delineated three classes of fetal cell binding antibodies based on the amino acid sequence of their heavy chains. Each class included subtypes identified first, by amino acid changes in and around the CDR3 region which reflect affinity maturation, and second, by association of different kappa chains.

Table 4 shows the distribution of different phage antibody types at different stages of enrichment on HEL or Fetal cells with or without preadsorption on adult cells. It is likely that the three classes of phage antibodies recognize three different epitopes based upon the difference in their staining profiles on fetal liver and adult cells.

After the fourth round of enrichment (in the enrichment series including preadsorption on adult cells) only phab type 5 was eluted from fetal liver cells. The selection for this type under the most stringent wash conditions suggests that it is the highest affinity combination of heavy and light chains.

Table 4

Phab Pool	Phage Antibody Type 1	2	3	4	5	6	7	U[a]
Primary	1							7
H3	1	1	2	1			1	
H4	3				1	1		
F3	10				6			
F4	10				6			
Hd3	5				1			
Hd4						16		
Fd1	8							
Fd2	6							
Fd3	12				1	20		1
Fd4	16							

[a] Unknown specificities

Mock enrichment experiments, in which three different phage antibodies spiked 1 in 10^6 non-specific control M13 phage were enriched on fetal liver cells (Table 5), demonstrated a dramatic 10^5-10^6 fold single round enrichment of these phage antibodies on fetal liver cells. This order of magnitude of enrichment is at the upper limits of those reported for enrichment of combinatorially-derived antibodies using purified antigen, which is of import when it is considered that the epitopes targeted in the present example are highly complex cell-surface antigens and have not been purified in any way.

Table 5
Enrichment of phage antibodies on fetal liver cells.

Phab	% Binding[b]	Starting Ratio[c]	Final Ratio[d]	Enrichment[e]
H3-3	1.6	1 in 4,500,000	1.6:1	3,600,000 fold
Fd3-1	1.1	1 in 150,000	1.1:1	165,000 fold
F4-7	0.5	1 in 100,000	3.0:1	300,000 fold

[b] Number of phabs eluted from cells after 10 washes, divided by the total phage loaded.
[c] Ratio of specific to non-specific phab in starting population.
[d] Ratio of specific to non-specific phab after elution from fetal cells.
[e] Final ratio divided by starting ratio.

The vast number of cell surface binding isolates seen on a consistent cell source (HEL) reflects the efficiency and diversity of the library constructed. The combination of tolerance immunization along with the efficiency of the present methods for library construction and amplification have yielded a remarkable library of phage antibodies which can be very rapidly enriched on whole cells. Such results are important for identifying phabs highly specific for a particular target cell-type from different individuals. This point is particularly emphasized by the fact that most of the phabs (even with tolerization immunization) isolated without preadsorption also recognize adult cells. In addition, enrichment at each stage with fetal liver cells from an independent fetus eliminates those antibodies which recognize individual specific markers. This added stringency in the enrichment for pan-fetal specific antibodies emphasizes the power of the present approach, which yielded 13 different versions of three classes of pan-fetal specific antibodies. In contrast, using the identical tolerization approach with conventional hybridoma methods, only two IgGs of the same type were obtained.

IV. Example 3: Affinity Determination of H3-3 and FB3-2 Antibodies

Affinity of purified antibodies was determined by Scatchard analysis. Varying amounts of antibody in significant excess were incubated for 16 hours at 4°C with a constant number of HEL cells. After extensive washes, bound antibody was eluted from cells at pH 2, and quatitated in an ELISA. For Scatchard analysis, free antibody was assumed to be equivalent to the total added. The K_a for each antibody was obtained from the negative slope of the line from the plot of bound versus bound/free antibody. All points were done in triplicate; the correlation coefficient for all reported slopes was greater than 90%.

V. Example 4: Specificity of H3-3 and FB3-2 Antibodies

In order to compare hybridoma-derived antibodies, such as anti-CD71 and anti-EM, with the subject antibodies, reactivity of these antibodies with fetal and maternal cells was tested by analytical flow cytometry (FACS efficiency assay). Briefly, $1x10^6$ cells per sample were stained with indicated amounts of FITC-conjugated pure antibody. 10,000 cells were analyzed for each sample. The results, provided in Table 3 above, are reported as "% positive", indicating the percentage of cells that were found to stain above background fluorescence as established by an isotype-matched negative control antibody.

The highly specific recognition of the H3-3 antibody for fetal as opposed to adult hematopoietic cells was further demonstrated by FACS and subsequent fluorescent *in situ* hybridization (FISH) analysis of sorted cells. Briefly, 400 fetal liver cells, demonstrated to

be male by Y-PCR, were spiked into one million adult PBMC. The spiked sample was then stained with Hoescht-DNA dye and 1μg of FITC-conjugated H3-3 antibody. The nucleated cells (those positive for Hoescht) were sorted for H3-3-positives, fixed to slides and analyzed for the presence of male cells by FISH. Male (Y) probe was detected with Cy3; female (X) probe by FITC. The results are summarized as follows: starting purity of fetal cells = 0.04%; background staining of adult cells = 0.05%; fetal (male) cells recovered = 301; purity of H3-3 sorted fetal cells = 36.4%.

All of the above-cited references and publications are hereby incorporated by reference.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, numerous equivalents to the specific method and reagents described herein. Such equivalents are considered to be within the scope of this invention and are covered by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Genzyme Corporation
(B) STREET: One Kendall Square
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 02139
(G) TELEPHONE: (508) 872-8400
(H) TELEFAX: (508) 872-5415

(ii) TITLE OF INVENTION: Process for Generating Specific Antibodies

(iii) NUMBER OF SEQUENCES: 61

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII (text)

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATATGCGGCC GCAGGTCTCC TCCTCTTAGC AGCACAACCA GCAATGGCCG ACATTSTGAT 60

GACDCAGTCT CCA 73

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATATGCGGCC GCAGGTCTCC TCCTCTTAGC AGCACAACCA GCAATGGCCG ATATCCAGAT 60

GACACAGACT HCA 73

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATATGCGGCC GCAGGTCTCC TCCTCTTAGC AGCACAACCA GCAATGGCCG ATGTTGTGMT 60

GACCCARACT CCA 73

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATATGCGGCC GCAGGTCTCC TCCTCTTAGC AGCACAACCA GCAATGGCCG ACATTGTGMT 60

GACMCAGWCT CCA 73

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATATGCGGCC GCAGGTCTCC TCCTCTTAGC AGCACAACCA GCAATGGCCC AAATTGTTCT 60

CACCCAGTCT CCA 73

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CATGGCCGGT TGGGCCGCGA GTAATAACAA TCCAGCGGCT GCCGTAGGCA ATAGGTATTT 60

CATTATGACT GTCTCCTTGC TATTAACACT CATTCCTGTT GAAGCTC 107

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTCGCTCGCC CATATGCGGC CGCAGGTCTC CTC 33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGGTTCGGC CCACATGGCC GGTTGGGCCG CGA 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCGCTCGCC CATCGCGGCC CAACCGGCCA TGG 33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTGGTTCGGC CCAAGGCTTA CTAGTACAAT CCC 33

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCGGCCCAAC CGGCCATGGC CG 22

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGGCCGCGAG TAATAACAAT CCAGCGGCTG CCGTAGGCAA TAGGTATTTC ATTATGACTG 60

TCTCCTTG 68

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCGCGGCCCA ACCGGCCATG GCCGAGGTCC AGCTKCAGCA GTCWG 45

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCGCGGCCCA ACCGGCCATG GCCGAGGTGA WGSTGGTGGA RTCTG 45

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TCGCGGCCCA ACCGGCCATG GCCCAGGTYC AGCTGMAGCA GTCTG 45

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGCGGCCCA ACCGGCCATG GCCGAGGTYC AGCTSCAGCA GTCTG 45

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCGCGGCCCA ACCGGCCATG GCCGAGGTGA AGCTKRTSGA GTCTG 45

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TCGCGGCCCA ACCGGCCATG GCCCAGGTGC AGCTKAAGSA GTCAG 45

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTGGTTCGGC CCAACTAGTA CAATCCCTGG GCACAATTTT C 41

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTGGTTCGGC CCAGATATCA CTAGTGGGCC CGCTGGGCTC AA 42

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTGGTTCGGC CCAACTAGTA GAACCTGGGG GGGTACTGG 39

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTGGTTCGGC CCATCTGCAC TAGTTGGAAT GGGCACATGC AG 42

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGGAATTCAT GGACTGGACC TGGAGGRTCY TCTKC 35

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGGAATTCAT GGAGYTTGGG CTGASCTGGS TTTT 34

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGGAATTCAT GRAMMWACTK TGKWSCWYSC TYCTG 35

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGGAATTCAT GGACATGRRR DYCCHVGYGT CASCTT 36

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGGAATTCAT GRCCTGSWCY CCTCTCYTYC TSWYC 35

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CCAAGCTTAG ACGAGGGGGA AAAGGGTT 28

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCAAGCTTGG AGGAGGGTGC CAGGGGG 27

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCAAGCTTGA AGCTCCTCAG AGGAGGG 27

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCAAGCTTTC ATCAGATGGC GGGAAGA 27

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Asp Pro Leu Tyr Gly Ser
1 5

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ser Gln Ser Thr His Val Leu Thr
1 5

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Ala Leu Lys Val His Met
1 5

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Asp Pro Leu Tyr Gly Asn
1 5

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gln Gln Trp Ser Ser Asn Pro Pro Thr
1 5

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Ser Gln Ser His His Val Leu Thr
1 5

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Asp Pro Leu Tyr Gly Asp
1 5

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Gly Asp Tyr Gly Asn Tyr Gly Asp Tyr Phe Asp Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Gln His Ser Trp Glu Ile Pro Tyr Thr

1 5

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Gln Asp Ser Trp Glu Ile Pro Tyr Thr
1 5

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Gln Gln Ser Asn Glu Asp Pro Tyr Thr
1 5

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Gln Gln Ser Asn Glu Asp Pro Phe Thr
1 5

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Gly Asp Tyr Gly Lys Tyr Gly Asp Tyr Phe Asp His
1 5 10

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Gly Val Tyr Gly Lys Tyr Gly Asp Tyr Phe Asp His
1 5 10

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Gln His Ser Trp Glu Ile Pro Phe Thr
1 5

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Cys Gly Gly Arg
1

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Glu Gly Tyr Gly Pro Thr Gly Tyr Tyr Ser Ala Met Asp Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Gln Gln Gly Tyr Ser Tyr Leu Thr
1 5

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 735 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..735

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```
ATGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCGGC CCAACCGGCC      60

ATGGCC GAG GTC CAG CTG CAG CAG TCT GGA CCT GAG CTG ATG ATG CCT      108
       Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Met Met Pro
         1               5                  10

GGG GCC TCA GTG AAG ATC TCC TGC AAG GCT ACT GGC TAC ACA TTG AGT      156
Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Thr Gly Tyr Thr Leu Ser
 15                  20                  25                  30

AGT TAC TGG CTA GAG TGG GTG AAA CAG AGC CCT GGA CAT GGC CTT GAA      204
Ser Tyr Trp Leu Glu Trp Val Lys Gln Ser Pro Gly His Gly Leu Glu
                 35                  40                  45

TGG ATC GGA GAG ATT TTA TTT GGA AGT GGT AGT GCT CAC TAC AAT GAG      252
Trp Ile Gly Glu Ile Leu Phe Gly Ser Gly Ser Ala His Tyr Asn Glu
             50                  55                  60

AAA TTC AAG GGC AAG GCC ACA TTC ACT GTA GAT ACA TCC TCC AAC ACA      300
Lys Phe Lys Gly Lys Ala Thr Phe Thr Val Asp Thr Ser Ser Asn Thr
         65                  70                  75

GCC TAC ATG CAA CTC AGC AGC CTG ACA TCT GAG GAC TCT GCC GTC TAT      348
Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr
     80                  85                  90

TAC TGT GCC AGA GGA GAC TAT GGT AAC TAC GGG GAC TAC TTT GAC TAC      396
Tyr Cys Ala Arg Gly Asp Tyr Gly Asn Tyr Gly Asp Tyr Phe Asp Tyr
 95                  100                 105                 110

TGG GGC CAA GGC ACC ACT CTC ACA GTC TCC TCA GCC AAA ACG ACA CCC      444
Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr Thr Pro
                115                 120                 125

CCA TCT GTC TAT CCA CTG GCC CCT GGA TCT GCT GCC CAA ACT AAC TCC      492
Pro Ser Val Tyr Pro Leu Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser
            130                 135                 140

ATG GTG ACC CTG GGA TGC CTG GTC AAG GGC TAT TTC CCT GAG CCA GTG      540
Met Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Pro Val
        145                 150                 155

ACA GTG ACC TGG AAC TCT GGA TCC CTG TCC AGC GGT GTG CAC ACC TTC      588
Thr Val Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe
    160                 165                 170

CCA GCT GTC CTG CAG TCT GAC CTC TAC ACT CTG AGC AGC TCA GTG ACT      636
Pro Ala Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr
```

```
175                 180                 185                 190

GTC CCC TCC AGC ACC TGG CCC AGC GAG ACC GTC ACC TGC AAC GTT GCC      684
Val Pro Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala
                195                 200                 205

CAC CCG GCC AGC AGC ACC AAG GTG GAC AAG AAA ATT GTG CCC AGG GAT      732
His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp
            210                 215                 220

TGT                                                                  735
Cys
```

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 223 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```
Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Met Met Pro Gly Ala
 1               5                  10                  15

Ser Val Lys Ile Ser Cys Lys Ala Thr Gly Tyr Thr Leu Ser Ser Tyr
            20                  25                  30

Trp Leu Glu Trp Val Lys Gln Ser Pro Gly His Gly Leu Glu Trp Ile
        35                  40                  45

Gly Glu Ile Leu Phe Gly Ser Gly Ser Ala His Tyr Asn Glu Lys Phe
    50                  55                  60

Lys Gly Lys Ala Thr Phe Thr Val Asp Thr Ser Ser Asn Thr Ala Tyr
65                  70                  75                  80

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
                85                  90                  95

Ala Arg Gly Asp Tyr Gly Asn Tyr Gly Asp Tyr Phe Asp Tyr Trp Gly
            100                 105                 110

Gln Gly Thr Thr Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Pro Ser
        115                 120                 125

Val Tyr Pro Leu Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser Met Val
    130                 135                 140

Thr Leu Gly Cys Leu Val Lys Gly Tyr Phe Pro Glu Pro Val Thr Val
145                 150                 155                 160

Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe Pro Ala
                165                 170                 175
```

Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr Val Pro
180 185 190

Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala His Pro
195 200 205

Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp Cys
210 215 220

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..399

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATGAAATACC TATTGCCTAC GGCGGCCGCA GGTCTCCTCC TCTTAGCAGC ACAACCAGCA 60

ATGGCC GAC ATT GTG ATG ACC CAG TCT CCT GCT TCC TTA GCT GTA TCT 108
Asp Ile Val Met Thr Gln Ser Pro Ala Ser Leu Ala Val Ser
1 5 10

CTG GGG CAG AGG GCC ACC ATC TCA TGC AGG GCC AGC CAA AGT GTC AGT 156
Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Gln Ser Val Ser
15 20 25 30

ACA TCT AGA TAT AGT TAT ATG CAC TGG TAC CAA CAG AAA CCA GGA CAG 204
Thr Ser Arg Tyr Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln
35 40 45

CCA GCC AAA CTC CTC ATC AAG TTT GCA TCC AAC CTA GAA TCT GGG GTC 252
Pro Ala Lys Leu Leu Ile Lys Phe Ala Ser Asn Leu Glu Ser Gly Val
50 55 60

CCT GCC AGG TTC AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC 300
Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn
65 70 75

ATC CAT CCT GTG GAG GAG GAG GAT ACT GCA ACA TAT TAC TGT CAG CAC 348
Ile His Pro Val Glu Glu Glu Asp Thr Ala Thr Tyr Tyr Cys Gln His
80 85 90

AGT TGG GAG ATT CCG TAC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATA 396
Ser Trp Glu Ile Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
95 100 105 110

AAA
Lys 399

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 111 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Asp Ile Val Met Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly
1 5 10 15

Gln Arg Ala Thr Ile Ser Cys Arg Ala Ser Gln Ser Val Ser Thr Ser
20 25 30

Arg Tyr Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln Pro Ala
35 40 45

Lys Leu Leu Ile Lys Phe Ala Ser Asn Leu Glu Ser Gly Val Pro Ala
50 55 60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His
65 70 75 80

Pro Val Glu Glu Glu Asp Thr Ala Thr Tyr Tyr Cys Gln His Ser Trp
85 90 95

Glu Ile Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys
100 105 110

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 735 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..735

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

ATGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCGGC CCAACCGGCC 60

ATGGCC GAG GTC CAG CTG CAG CAG TCT GGA GCT GAG CTG ATG ATG CCT 108

```
          Glu Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Met Met Pro
            1               5                  10

GGG GCC TCA GTG AAG ATC TCC TGC AAG GCT ACT GGC TAC ACA TTG AGT      156
Gly Ala Ser Val Lys Ile Ser Cys Lys Ala Thr Gly Tyr Thr Leu Ser
 15                  20                  25                  30

AGT TCC TGG CTA GAG TGG GTG AAA CAG AGC CCT GGA CAT GGC CTT GAA      204
Ser Ser Trp Leu Glu Trp Val Lys Gln Ser Pro Gly His Gly Leu Glu
                 35                  40                  45

TGG ATT GGA GAG ATT TTA TTT GGA AGT GGT AGT GCT CAC TAC AAT GAG      252
Trp Ile Gly Glu Ile Leu Phe Gly Ser Gly Ser Ala His Tyr Asn Glu
             50                  55                  60

AAA TTC AAG GGC AAG GCC ACA TTC ACT GTA GAT ACA TCC TCC AAC ACA      300
Lys Phe Lys Gly Lys Ala Thr Phe Thr Val Asp Thr Ser Ser Asn Thr
         65                  70                  75

GCC TAC ATG CAA CTC AGC AGC CTG ACA TCT GAG GAC TCT GCC GTC TAT      348
Ala Tyr Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr
     80                  85                  90

TAC TGT GCC AGA GGA GAC TAT GGT AAC TAC GGG GAC TAC TTT GAC TAC      396
Tyr Cys Ala Arg Gly Asp Tyr Gly Asn Tyr Gly Asp Tyr Phe Asp Tyr
 95                  100                 105                 110

TGG GGC CAA GGC CAA GCT CTC ACA GTC TTC TCA GCC AAA ACG ACA CCC      444
Trp Gly Gln Gly Gln Ala Leu Thr Val Phe Ser Ala Lys Thr Thr Pro
                115                 120                 125

TCA TCT GTC TAT CCA CTG GCT GCT GGA TCT GCT GCC CAA ACT AAC TCC      492
Ser Ser Val Tyr Pro Leu Ala Ala Gly Ser Ala Ala Gln Thr Asn Ser
            130                 135                 140

ATG GTG ACC CTG GGA TGC CTG GTC AAG GGC TAT CTC CCT GAG CCA GTG      540
Met Val Thr Leu Gly Cys Leu Val Lys Gly Tyr Leu Pro Glu Pro Val
        145                 150                 155

ACA GTG ACC TGG AAC TCT GGA TCC CTG TCC AGC GGT GTG CAC ACC TTC      588
Thr Val Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe
    160                 165                 170

CCA GCT GTC CTG CAG TCT GAC CTC TAC ACT CTG AGC AGA TCA GTG ACT      636
Pro Ala Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Arg Ser Val Thr
175                 180                 185                 190

GTC CCC TCC AGC ACC TGG CCC AGC GAG ACC GTC ACC TGC AAC GTT GCC      684
Val Pro Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala
                195                 200                 205

CAC CCG GCC AGC AGC ACC AAG GTG GAC AAG AAA ATT GTG CCC AGG GAT      732
His Pro Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp
            210                 215                 220

TGT                                                                  735
Cys
```

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 223 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
Glu Val Gln Leu Gln Gln Ser Gly Ala Glu Leu Met Met Pro Gly Ala
 1               5                  10                  15

Ser Val Lys Ile Ser Cys Lys Ala Thr Gly Tyr Thr Leu Ser Ser Ser
            20                  25                  30

Trp Leu Glu Trp Val Lys Gln Ser Pro Gly His Gly Leu Glu Trp Ile
        35                  40                  45

Gly Glu Ile Leu Phe Gly Ser Gly Ser Ala His Tyr Asn Glu Lys Phe
    50                  55                  60

Lys Gly Lys Ala Thr Phe Thr Val Asp Thr Ser Ser Asn Thr Ala Tyr
65                  70                  75                  80

Met Gln Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys
                85                  90                  95

Ala Arg Gly Asp Tyr Gly Asn Tyr Gly Asp Tyr Phe Asp Tyr Trp Gly
            100                 105                 110

Gln Gly Gln Ala Leu Thr Val Phe Ser Ala Lys Thr Thr Pro Ser Ser
        115                 120                 125

Val Tyr Pro Leu Ala Ala Gly Ser Ala Ala Gln Thr Asn Ser Met Val
    130                 135                 140

Thr Leu Gly Cys Leu Val Lys Gly Tyr Leu Pro Glu Pro Val Thr Val
145                 150                 155                 160

Thr Trp Asn Ser Gly Ser Leu Ser Ser Gly Val His Thr Phe Pro Ala
                165                 170                 175

Val Leu Gln Ser Asp Leu Tyr Thr Leu Ser Arg Ser Val Thr Val Pro
            180                 185                 190

Ser Ser Thr Trp Pro Ser Glu Thr Val Thr Cys Asn Val Ala His Pro
        195                 200                 205

Ala Ser Ser Thr Lys Val Asp Lys Lys Ile Val Pro Arg Asp Cys
    210                 215                 220
```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..720

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
ATGAAATACC TATTGCCTAC GGCGGCCGCA GGTCTCCTCC TCTTAGCAGC ACAACCAGCA      60

ATGGCC GAC ATT GTG ATG ACC CAG TCT CCT GCT TCC TTA GCT GTA TCT          108
       Asp Ile Val Met Thr Gln Ser Pro Ala Ser Leu Ala Val Ser
         1               5                  10

CTG GGG CAG AGG GCC ACC ATC TCA TGC AGG GTC AGG CAA AGT GTC AGT        156
Leu Gly Gln Arg Ala Thr Ile Ser Cys Arg Val Arg Gln Ser Val Ser
 15                  20                  25                  30

ACA TCT AGC CAT AGT TAT ATG CAC TGG TAC CAA CAG AAA CCA GGA CAG        204
Thr Ser Ser His Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln
                35                  40                  45

CCA CCC AAA CTC CTC ATC AAG TAT GCA TCC AAC CTA GAA TCT GGG GTC        252
Pro Pro Lys Leu Leu Ile Lys Tyr Ala Ser Asn Leu Glu Ser Gly Val
            50                  55                  60

CCT GCC AGG TTC AGT GGC AGT GGG TCT GGG ACA GAC TTC ACC CTC AAC        300
Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn
        65                  70                  75

ATC CAT CCT GTG GAG GAG GAG GAT ACT GCA ACA TAT TAC TGT CAG CAC        348
Ile His Pro Val Glu Glu Glu Asp Thr Ala Thr Tyr Tyr Cys Gln His
    80                  85                  90

AGT TGG GAG ATT CCG TAC ACG TTC GGA GGG GGG ACC AAG CTG GAA ATA        396
Ser Trp Glu Ile Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile
 95                 100                 105                 110

AAA CGG GCT GAT GCT GCA CCA ACT GTA TCC ATC TTC CCA CCA TCC AGT        444
Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser
                115                 120                 125

GAG CAG TTA ACA TCT GGA GGT GCC TCA GTC GTG TGC TTC TTG AAC AAC        492
Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe Leu Asn Asn
            130                 135                 140

TTC TAC CCC AAA GAC ATC AAT GTC AAG TGG AAG ATT GAT GGC AGT GAA        540
Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu
        145                 150                 155

CGA CAA AAT GGC GTC CTG AAC AGT TGG ACT GAT CAG GAC AGC AAA GAC        588
```

```
Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp
    160                 165                 170

AGC ACC TAC AGC AGG AGC AGC ACC CTC ACG TTG ACC AAG GAC GAG TAT      636
Ser Thr Tyr Ser Arg Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr
175                 180                 185                 190

GAA CGA CAT AAC AGC TAT ACC TGT GAG GCC ACT CAC AAG ACA TCA ACT      684
Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr
                195                 200                 205

TCA CCC ATT GTC AAG AGC TTC AAC AGG AAT GAG TGT TAA                  723
Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
            210                 215
```

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

```
Asp Ile Val Met Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly
  1               5                  10                  15

Gln Arg Ala Thr Ile Ser Cys Arg Val Arg Gln Ser Val Ser Thr Ser
            20                  25                  30

Ser His Ser Tyr Met His Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro
        35                  40                  45

Lys Leu Leu Ile Lys Tyr Ala Ser Asn Leu Glu Ser Gly Val Pro Ala
    50                  55                  60

Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Asn Ile His
65                  70                  75                  80

Pro Val Glu Glu Glu Asp Thr Ala Thr Tyr Tyr Cys Gln His Ser Trp
                85                  90                  95

Glu Ile Pro Tyr Thr Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys Arg
            100                 105                 110

Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser Glu Gln
        115                 120                 125

Leu Thr Ser Gly Gly Ala Ser Val Val Cys Phe Leu Asn Asn Phe Tyr
    130                 135                 140

Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu Arg Gln
145                 150                 155                 160

Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp Ser Thr
```

```
                165                 170                 175

Tyr Ser Arg Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr Glu Arg
            180                 185                 190

His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr Ser Pro
        195                 200                 205

Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
    210                 215
```

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 717 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..717

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
GTGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCGGC CCAACCGGCC      60

ATGGCC GAG GTG AAG CTT ATG GAG TCT GGG GGA GAC TTA GTG AAG CCT      108
       Glu Val Lys Leu Met Glu Ser Gly Gly Asp Leu Val Lys Pro
         1               5                  10

GGA GGG TCC CTG AAA CTC TCC TGT GCA GCC TCT GGA TTC ACT TTC AGT      156
Gly Gly Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
 15                  20                  25                  30

GAC TAT TAC ATG TAT TGG GTT CGC CAG ACT CCG GAA AAG AGG CTG GAG      204
Asp Tyr Tyr Met Tyr Trp Val Arg Gln Thr Pro Glu Lys Arg Leu Glu
                 35                  40                  45

TGG GTC GCA ACC ATT AGT GAT GAT GGT ACT TAC ACC TAC TAT GCA GAC      252
Trp Val Ala Thr Ile Ser Asp Asp Gly Thr Tyr Thr Tyr Tyr Ala Asp
             50                  55                  60

AGT GTG AAG GGG CGA TTC ACC ATC TCC AGA GAC AAT GCC AAG AAC AAC      300
Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Asn
         65                  70                  75

CTC TAC CTG CAA ATG AAC AGT CTG AAG TCT GAG GAC ACA GCC ATG TAT      348
Leu Tyr Leu Gln Met Asn Ser Leu Lys Ser Glu Asp Thr Ala Met Tyr
     80                  85                  90

TAC TGT GCA AGA GAT CCC CTT TAT GGC AGC TGG GGC CAA GGC ACC ACT      396
Tyr Cys Ala Arg Asp Pro Leu Tyr Gly Ser Trp Gly Gln Gly Thr Thr
 95                  100                 105                 110
```

```
CTC ACA GTC TCC TCA GCC AAA ACG ACA CCC CCA TCT GTC TAT CCA CTG        444
Leu Thr Val Ser Ser Ala Lys Thr Thr Pro Pro Ser Val Tyr Pro Leu
                115                 120                 125

GCC CCT GGA TCT GCT GCC CAA ACT AAC TCC ATG GTG ACC CTG GGA TGC        492
Ala Pro Gly Ser Ala Ala Gln Thr Asn Ser Met Val Thr Leu Gly Cys
            130                 135                 140

CTG GTC AAG GGC TAT TTC CCT GAG CCA GTG ACA GTG ACC TGG AAC TCT        540
Leu Val Lys Gly Tyr Phe Pro Glu Pro Val Thr Val Thr Trp Asn Ser
        145                 150                 155

GGA TCC CTG TCC AGC GGT GTG CAC ACC TTC CCA GCT GTC CTG CAG TCT        588
Gly Ser Leu Ser Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser
    160                 165                 170

GAC CTC TAC ACT TTG AGC AGC TCA GTG ACT GTC CCC TCC AGC ACC TGG        636
Asp Leu Tyr Thr Leu Ser Ser Ser Val Thr Val Pro Ser Ser Thr Trp
175                 180                 185                 190

TCC AGC GAG ACC GTC ACC TGC AAC GTT GCC CAC CCG GCC AGC AGC ACC        684
Ser Ser Glu Thr Val Thr Cys Asn Val Ala His Pro Ala Ser Ser Thr
                195                 200                 205

AAG GTG GAC AAG AAA ATT GTG CCC AGG GAT TGT                            717
Lys Val Asp Lys Lys Ile Val Pro Arg Asp Cys
            210                 215
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 217 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
Glu Val Lys Leu Met Glu Ser Gly Gly Asp Leu Val Lys Pro Gly Gly
  1               5                  10                  15

Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asp Tyr
            20                  25                  30

Tyr Met Tyr Trp Val Arg Gln Thr Pro Glu Lys Arg Leu Glu Trp Val
        35                  40                  45

Ala Thr Ile Ser Asp Asp Gly Thr Tyr Thr Tyr Tyr Ala Asp Ser Val
    50                  55                  60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Asn Leu Tyr
65                  70                  75                  80

Leu Gln Met Asn Ser Leu Lys Ser Glu Asp Thr Ala Met Tyr Tyr Cys
                85                  90                  95
```

```
Ala Arg Asp Pro Leu Tyr Gly Ser Trp Gly Gln Gly Thr Thr Leu Thr
            100                 105                 110

Val Ser Ser Ala Lys Thr Thr Pro Pro Ser Val Tyr Pro Leu Ala Pro
        115                 120                 125

Gly Ser Ala Ala Gln Thr Asn Ser Met Val Thr Leu Gly Cys Leu Val
    130                 135                 140

Lys Gly Tyr Phe Pro Glu Pro Val Thr Val Thr Trp Asn Ser Gly Ser
145                 150                 155                 160

Leu Ser Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Asp Leu
                165                 170                 175

Tyr Thr Leu Ser Ser Ser Val Thr Val Pro Ser Ser Thr Trp Ser Ser
            180                 185                 190

Glu Thr Val Thr Cys Asn Val Ala His Pro Ala Ser Ser Thr Lys Val
        195                 200                 205

Asp Lys Lys Ile Val Pro Arg Asp Cys
    210                 215
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 67..720

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
ATGAAATACC TATTGCCTAC GGCGGCCGCA GGTCTCCTCC TCTTAGCAGC ACAACCAGCA      60

ATGGCC GAT GTT GTG CTG ACC CAG ACT CCA CTC TCC CTG CCT GTC AGT          108
       Asp Val Val Leu Thr Gln Thr Pro Leu Ser Leu Pro Val Ser
         1               5                  10

CTT GGA GGT CAA GCC TCC ATC TCT TGC AGA TCT AGT CAG AGC CTT GTA         156
Leu Gly Gly Gln Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Val
 15                  20                  25                  30

CAC AGT AAT GGA AAC ACC TAT TTA CAT TGG TAC CTG CAG AAG CCA GGC         204
His Ser Asn Gly Asn Thr Tyr Leu His Trp Tyr Leu Gln Lys Pro Gly
                 35                  40                  45

CAG TCT CCA AAG CTC CTG ATC TAC AAG GTT TCC AAC CGG TTT TCT GGG         252
```

```
Gln Ser Pro Lys Leu Leu Ile Tyr Lys Val Ser Asn Arg Phe Ser Gly
            50                  55                  60

GTC CCA GAC AGG TTC AGT GGC AGT GGA TCA GGG ACA GAT TTC ACA CTC      300
Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu
         65                  70                  75

AAG ATC AGC AGA GTG GAG GCT GAG GAT CTG GGA GTT TAT TTC TGC TCT      348
Lys Ile Ser Arg Val Glu Ala Glu Asp Leu Gly Val Tyr Phe Cys Ser
      80                  85                  90

CAA AGT ACA CAT GTT CTC ACG TTC GGT GCT GGG ACC AAG CTG GAG CTG      396
Gln Ser Thr His Val Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu
 95                 100                 105                 110

AAA CGG GCT GAT GCT GCA CCA ACT GTA TCC ATC TTC CCA CCA TCC AGT      444
Lys Arg Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser
                115                 120                 125

GAG CAG TTA ACA TCT GGA GGT GCC TCA GTC GTG GGC TTC TTG AAC AAC      492
Glu Gln Leu Thr Ser Gly Gly Ala Ser Val Val Gly Phe Leu Asn Asn
            130                 135                 140

TTC TAC CCC AAA GAC ATC AAT GTC AAG TGG AAG ATT GAT GGC AGT GAA      540
Phe Tyr Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu
        145                 150                 155

CGA CAA AAT GGC GTC CTG AAC AGT TGG ACT GAT CAG GAC AGC AAA GAC      588
Arg Gln Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp
    160                 165                 170

AGC ACC TAC AGC AGG AGC AGC ACC CTC ACG TTG ACC AAG GAC GAG TAT      636
Ser Thr Tyr Ser Arg Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr
175                 180                 185                 190

GAA CGA CAT AAC AGC TAT ACC TGT GAG GCC ACT CAC AAG ACA TCA ACT      684
Glu Arg His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr
                195                 200                 205

TCA CCC ATT GTC AAG AGC TTC AAC AGG AAT GAG TGT TAA                  723
Ser Pro Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
            210                 215
```

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 218 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Asp Val Val Leu Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Leu Gly
 1               5                  10                  15
```

```
Gly Gln Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Val His Ser
            20                  25                  30

Asn Gly Asn Thr Tyr Leu His Trp Tyr Leu Gln Lys Pro Gly Gln Ser
        35                  40                  45

Pro Lys Leu Leu Ile Tyr Lys Val Ser Asn Arg Phe Ser Gly Val Pro
    50                  55                  60

Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65                  70                  75                  80

Ser Arg Val Glu Ala Glu Asp Leu Gly Val Tyr Phe Cys Ser Gln Ser
                85                  90                  95

Thr His Val Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Arg
            100                 105                 110

Ala Asp Ala Ala Pro Thr Val Ser Ile Phe Pro Pro Ser Ser Glu Gln
        115                 120                 125

Leu Thr Ser Gly Gly Ala Ser Val Val Gly Phe Leu Asn Asn Phe Tyr
    130                 135                 140

Pro Lys Asp Ile Asn Val Lys Trp Lys Ile Asp Gly Ser Glu Arg Gln
145                 150                 155                 160

Asn Gly Val Leu Asn Ser Trp Thr Asp Gln Asp Ser Lys Asp Ser Thr
                165                 170                 175

Tyr Ser Arg Ser Ser Thr Leu Thr Leu Thr Lys Asp Glu Tyr Glu Arg
            180                 185                 190

His Asn Ser Tyr Thr Cys Glu Ala Thr His Lys Thr Ser Thr Ser Pro
        195                 200                 205

Ile Val Lys Ser Phe Asn Arg Asn Glu Cys
    210                 215
```

CLAIMS:

1. A method for generating a specific antibody for an immunorecessive epitope, and nucleic acid encoding said antibody, comprising the steps of

generating an immunotolerance-derived antibody repertoire for an immunorecessive epitope;

generating an antibody display library comprising a variegated V-gene library expressed by a population of display packages, said V-gene library cloned from said antibody repertoire, and

selecting display pakages of said antibody display library which have a desired binding specificity for said immunorecessive epitope.

2. A method for generating a specific antibody for immunorecessive epitope, and nucleic acid encoding said antibody, comprising the steps of

generating a variegated display library of antibody variable regions, said antibody variable regions cloned from an immunotolerance-derived antibody repertoire, and

selecting antibody variable regions of said display library which have a desired binding specificity for an immunorecessive epitope.

3. The method of claim 2, wherein said display library is a phage display library.

4. The method of claim 2, wherein said display library is a bacterial cell-surface display library or a spore display library.

5. The method of claim 2, wherein said antibody variable region is a heavy chain, a light chain, a heavy-light chain pair, a V_H, a V_L, an Fab, an Fd, an Fv, or an scFv.

6. The method of claim 2, wherein said immunorecessive epitope is a cell-type specific marker.

7. The method of claim 6, wherein said immunorecessive epitope is a cancer cell marker.

8. The method of claim 6, wherein said immunorecessive epitope is a fetal cell marker.

9. The method of claim 6, wherein said immunorecessive epitope is a stem cell marker.

10. The method of claim 2, wherein said immunorecessive epitope comprises at least one amino acid residue in a variant protein that is different from a related or parent protein.

11. The method of claim 2, wherein said immunotolerance-derived antibody repertoire is generated by chemical immunosuppression.

12. The method of claim 11, wherein said immunotolerance-derived antibody repertoire is generated by cyclophosphamide-induced immunosuppression.

13. The method of claim 2, wherein said immunotolerance-derived antibody repertoire is generated by neonatal tolerization.

14. The method of claim 2, wherein said antibody variable regions are selected from said display library by a differential binding means comprising affinity separation of antibody variable regions which specifically bind said epitope from antibody variable regions which do not specifically bind.

15. The method of claim 14, wherein said differential binding means comprises panning said display library on a cell surface comprising said epitope.

16. A method for generating a specific antibody to an immunorecessive epitope, and genes encoding said antibody, comprising:

(a) transforming suitable host cells with a library of replicable phage vectors encoding a library of phage particles displaying a fusion antibody/coat protein, said fusion protein comprising a phage coat protein portion and an antibody variable region portion, said antibody variable region portion being obtained from an immunotolerance-derived variegated V-gene library;

(b) culturing said transformed host cells such that said phage particles are formed and said fusion protein are expressed; and

(c) selecting any of said phage particles having an antibody variable region portion which binds to a an immunorecessive epitope.

17. The method of claim 16, wherein said transformed host cells further comprise a second antibody gene encoding a second variable region which is expressed in said transformed host cells, and which associates with said antibody variable region portion of said fusion protein to form a heterodimeric Fv.

18. The method of claim 17, wherein said second antibody gene is obtained from said V-gene library .

19. The method of claim 17, wherein said second variable region is a polypeptide chain apart from said fusion protein and further comprises a secretion signal sequence that enables said second variable region to be secreted from said transformed host cells.

20. The method of claim 17, wherein said phage vector further comprises said second antibody gene.

21. The method of claim 20, wherein said fusion protein further comprises said second variable region covalently linked to said antibody variable region portion to form a single polypeptide chain antibody.

22. The method of claim 16, wherein said phage particle is selected from a group consisting of M13, f1, fd, If1, Ike, Xf, Pf1, Pf3, λ ,T4, T7, P2, P4, ϕX-174, MS2 and f2.

23. The method of claim 16, wherein said phage particle is a filamentous bacteriophage specific for *Escherichia coli* and said phage coat protein is coat protein III.

24. The method of claim 23, wherein said filamentous bacteriophage is selected from a group consisting of M13, fd, and f1.

25. The method of claim 16, wherein said transformed host cells are cultured with a helper phage suitable for inducing formation of said phage particles.

26. The method of claim 16, wherein said phage particles are selected by a differential binding means comprising contacting said phage particles with said immunorecessive epitope and separating phage particles which specifically bind said epitope from phage particles which do not specifically bind said epitope.

27. The method of claim 26, wherein said differential binding means comprises an affinity chromatographic means in which said immunorecessive epitope is provided as a component of an insoluble matrix.

28. The method of claim 27, wherein said insoluble matrix comprises said immunorecessive epitope attached to a polymeric support.

29. The method of claim 27, wherein said insoluble matrix comprises a immunorecessive cell displaying said target epitope.

30. The method of claim 26, wherein said differential binding means comprises immunoprecipitating said phage particles with a multivalent form of said immunorecessive epitope, and subsequently removing non-specifically bound phage particles from said precipitate.

31. The method of claim 16, wherein said immunorecessive epitope is a cell-type specific marker.

32. The method of claim 31, wherein said immunorecessive epitope is a cancer cell marker.

33. The method of claim 31, wherein said immunorecessive epitope is a fetal cell marker.

34. The method of claim 31, wherein said immunorecessive epitope is a stem cell marker.

35. The method of claim 16, wherein said immunorecessive epitope comprises at least one amino acid residue in a variant protein that is different from a related or parent protein.

36. A method for generating a specific antibody for an immunorecessive epitope, and genes encoding said antibody, comprising:

(a) generating an immunotolerance-derived population of antibody-producing cells enriched for cells producing antibodies to an immunorecessive epitope;

(b) generating a variegated V-gene library encoding at least a variable region of immunoglobulin chains expressed by said enriched population of antibody-producing cells;

(c) generating a library of replicable phage vectors encoding a library of phage particles displaying a fusion coat protein, each of said phage vectors comprising a chimeric coat protein gene encoding said fusion coat protein, said chimeric gene including

(i) a first antibody gene encoding a variable region derived from said V-gene library, and

(ii) a second gene encoding at least a portion of a phage coat protein, such that said library of phage vectors encodes a plurality of cloned variable regions;

(d) transforming suitable host cells with said library of replicable phage vectors;

(e) culturing said transformed host cells such that said phage particles are formed and said fusion coat protein are expressed; and

(f) selecting any of said phage vectors corresponding to phage particles which display a variable region which binds to said immunorecessive epitope.

37. The method of claim 36, wherein said immunotolerance-derived population of antibody-producing cells are generated by chemical immunosuppressionof antibody production to immunodominant epitopes normally associated with said immunorecessive epitope in an immunogen.

38. The method of claim 36, wherein said immunotolerance-derived population of antibody-producing cells are generated by neonatal tolerization to suppress production of antibodies directed to immunodominant epitopes normally associated with said immunorecessive epitope in an immunogen.

39. A method for generating a specific antibody for a fetal cell-specific antigen, and nucleic acid encoding said antibody, comprising the steps of

generating a variegated display library of antibody variable regions, said antibody variable regions cloned from an immunotolerance-derived antibody repertoire enriched for antibodies to a fetal cell-specific antigen, and

selecting antibody variable regions of said display library which have a desired binding specificity for said fetal cell-specific antigen.

40. The method of claim 39, wherein said antibody variable regions of said display library are separated by a step comprising panning said display library on a fetal cell comprising said fetal cell-specific antigen.

41. The method of claim 39, wherein said fetal-cell specific antigen is a marker for fetal nucleated red blood cells.

42. An antibody that specifically binds an onco/fetal antigen said antibody having a heavy chain variable region comprising a CDR3 amino acid sequence selected from the group consisting of DPLYGS, DPLYGN, DPLYGD, GDYGDYGDYFDY, GDYGNYGDYFDY, GDYGKYGDYFDH, GVYGKYGDYFDH, and EGYGPTGYYSAMDY.

43. The antibody of claim 42, further comprising a light chain variable region comprising a CDR3 amino acid sequence selected from the group consisting of SQSTHVLT, ALKVHM, HSWEIPYT, QQWSSNPPT, SQSHHVLT, QHSWEIPYT, QDSWEIPYT, QQSNEDPYT, QQSNEDPFT, QQWSSNPPT, QHSWEIPFT, and GQGYSYLT.

44. An antibody isolated by the method of claim 1.

45. An antibody isolated by the method of claim 16.

46. An antibody isolated by the method of claim 36.

47. An antibody isolated by the method of claim 39.

48. An antibody display library enriched for specific antibodies to an immunorecessive epitope comprising, a variegated V-gene library expressed by a population of display packages and enriched for specific antibodies by differential binding with an immunorecessive epitope, said V-gene library cloned from an immunotolerance-derived antibody repertoire.

49. The antibody display library of claim 48, wherein said display package is a phage particle.

50. The antibody display library of claim 48, wherein said immunotolerance-derived antibody repertoire is generated with a set of immunogen and toleragen in which said immunorecessive epitope comprises a cell-type specific marker.

51. The antibody display library of claim 50, wherein said cell-type specific marker is a fetal nucleated red blood cell marker, and said toleragen comprises a maternal erythroid cell and said immunogen comprises a fetal erythroid cell.

52. The antibody display library of claim 50, wherein said cell-type specific marker is a tumor cell marker.

53. The antibody display library of claim 52, wherein said tumor cell marker is a colon cancer marker, and said toleragen comprises a normal colon cell and said immunogen comprises a colon carcinoma cell.

54. The antibody display library of claim 52, wherein said tumor cell marker is a metastatic tumor cell marker, and said toleragen comprises a non-metastatic tumor cell and said immunogen comprises a metastatic tumor cell.

55. The antibody display library of claim 50, wherein said cell-type specific marker is a precursor nerve cell marker, and said toleragen comprises a differentiated nerve cell and said immunogen comprises an embryonic nerve cell.

56. The antibody display library of claim 50, wherein said cell-type specific marker is a hematopoeitic cell marker, and said toleragen comprises a committed stem cell and said immunogen comprises a hematopoietic stem cell.

57. The antibody display library of claim 48, wherein said immunotolerance-derived antibody repertoire is generated with a set of immunogen and toleragen in which said immunorecessive epitope comprises a determinant unique to a variant form of a protein.

58. The antibody display library of claim 57, wherein said variant protein is Apolipoprotein E4, and said toleragen comprises a Apolipoprotein E and said immunogen comprises Apolipoprotein E4.

59. The antibody display library of claim 57, wherein said variant protein is a p53 mutant having one or more amino residues different from wild-type p53, and said toleragen comprises a wild-type p53 and said immunogen comprises said p53 mutant.

60. The antibody display library of claim 57, wherein said variant protein is a *ras* mutant having one or more amino residues different from wild-type *ras*, and said toleragen comprises a wild-type *ras* and said immunogen comprises said *ras* mutant.

61. A variegated population of antibodies cloned from the antibody display library of claim 48.

62. An isolated antibody of the antibody display library of claim 48.

63. (New) A method for generating an antibody having a binding association constant for an immunorecessive epitope of greater than $1x10^8M^{-1}$, and nucleic acid encoding said antibody, comprising the steps of

generating an immunotolerance-derived antibody repertoire for an immunorecessive epitope;

generating an antibody display library comprising a variegated V-gene library expressed by a population of display packages, said V-gene library cloned from said antibody repertoire, and

selecting display packages of said antibody display library which have a binding association constant for said immunorecessive epitope of greater than $1x10^8M^{-1}$.

64. (New) A method for generating an antibody selective for an immunorecessive epitope, and nucleic acid encoding said antibody, comprising the steps of

generating an immunotolerance-derived antibody repertoire for an immunorecessive epitope;

generating an antibody display library comprising a variegated V-gene library expressed by a population of display packages, said V-gene library cloned from said antibody repertoire, and

selecting display packages of said antibody display library which have a binding association constant for said immunorecessive epitope of greater than $1x10^8M^{-1}$ and a relative specificity of at least 10 fold over binding to background antigens.

65. (New) A method for generating an antibody which selectively binds an immunorecessive epitope unique to a first cell phenotype of a related population of cells, and nucleic acid encoding said antibody, comprising the steps of

generating an immunotolerance-derived antibody repertoire for an immunorecessive epitope on said first cell phenotype;

generating an antibody display library comprising a variegated V-gene library expressed by a population of display packages, said V-gene library cloned from said antibody repertoire;

generating an enriched display library by one or more of the steps of

(i) removing from said antibody display library those display packages with substantial background binding to cells of said related cell population other than said first cell phenotype, and

(ii) removing from said antibody display library those display packages which bind to said first cell phenotype in an individually selective manner,

said enriched display library comprising remaining display packages of said antibody display library; and

selecting display packages of said enriched display library which have a desired binding affinity for said first cell phenotype.

66. (New) The method of claim 65, wherein said immunorecessive epitope is a fetal cell marker.

67. (New) The method of claim 65, wherein said immunorecessive epitope is a cancer cell marker.

68. (New) The method of claim 65, wherein said immunorecessive epitope is a stem cell marker.

69. (New) The method of claim 65, wherein said display packages of said enriched display library are selected by panning said display packages will cells of said first cell phenotype.

70. (New) An antibody immunoreactive with a fetal cell surface antigen, said antibody having a binding association constant for said antigen in excess of $1x10^8$ M^{-1} and having no substantial background binding to maternal cells.

71. (New) An antibody specifically immunoreactive with a fetal cell surface antigen and characterized by a specificity of at least 10 fold over background binding to maternal antibodies.

72. (New) An antibody specifically immunoreactive with a fetal cell surface antigen, said antibody having a background binding to maternal cells of at least 2 fold less than an anti-CD71 antibody selected from the group consisting of a 5E9 antibody, an L5.1 antibody, and an L01.1 antibody.

73. (New) An antibody that binds an onco/fetal antigen, which antibody includes an antigen binding site comprising one or both of a first variable region and a second variable region, each of said first and second variable regions including complementarity determining regions of an H3-3 antibody, an FB3-2 antibody or an F4-7 antibody.

74. (New) The antibody of claim 73, wherein each of the first and second variable regions are represented by the general formula

FR(1)-CDR(1)-FR(2)-CDR(2)-FR(3)-CDR(3)-FR(4)

wherein FR(1)-FR(4) represent polypeptides from antibody framework regions, and CDR(1)-CDR(3) represent polypeptides from complementarity determining regions of an H3-3 antibody, an FB3-2 antibody or an F4-7 antibody.

75. (New) The antibody of claim 74, wherein each of the CDR(1), CDR(2), and CDR(3) for a single variable region have amino acid sequences selected from the group consisting:

CDR(1) = SYWLE, CDR(2) = EILFGSGSAHYNEKFKG, CDR(3) = GDYGNYGDYFDY;
CDR(1) = RASQSVSTSRYSYMH, CDR(2) = FASNLES, CDR(3) = HSWEIPYT;
CDR(1) = SSWLE, CDR(2) = EILFGSGSAHYNEKFKG, CDR(3) = GDYGNYGDYFDY;
CDR(1) = RVRQSVSTSSHSYMH, CDR(2) = YASNLES, CDR(3) = HSWEIPYT;
CDR(1) = DYYMY, CDR(2) = TISDDGTYTYYADSVKG, CDR(3) = DPLYGS; and
CDR(1) = RSSQSLVHSNGNTYLH, CDR(2) = KVSNRFS, CDR(3) = SQSTHVLT.

76. (New) The antibody of claim 73, wherein the variable regions are selected from group consisting of E1-S121 of SEQ ID No. 51, D1-K111 of SEQ ID No. 53, E1-S121 of SEQ ID No. 55, D1-K111 of SEQ ID No. 57, E1-S115 of SEQ ID No. 59, and D1-K111 of SEQ ID No. 61.

77. (New) The antibody of claim 73, which antibody further comprises framework region polypeptides from a human antibody.

78. (New) The antibody of claim 73, which antibody further comprises a constant region polypeptide from a human antibody.

79. (New) An antibody display library enriched for antibodies having binding constants for a cell surface antigen greater than $10^8 M^{-1}$, which antibody library comprises a variegated V-gene library expressed by a population of display packages and enriched for specific antibodies by differential binding with an immunorecessive epitope of said cell surface antigen, said V-gene library cloned from an immunotolerance-derived antibody repertoire.

80. (New) The antibody display library of claim 79 wherein said cell surface antigen is a fetal nucleated red blood cell marker.

81. (New) The antibody display library of claim 79, wherein said cell surface antigen is a tumor cell marker.

82. (New) A library of isolated nucleic acids encoding antigen binding sites immunoreactive with an immunorecessive epitope, comprising a variegated V-gene library encoding at least a variable region of immunoglobulin chains expressed by antibody-producing

cells of an animal, which antibody-producing cells are enriched by immunotolerization for cells producing antibodies to the immunorecessive epitope.

83. (New) The gene library of claim 82, wherein said V-gene library is expressed by a population of display packages.

84. (New) The gene library of claim 83, wherein said display package is a phage particle.

85. (New) The gene library of claim 82, wherein said immunorecessive epitope comprises an onco/fetal cell surface marker.

Figure 1A

primer number	Heavy Chain 5' Primer	SfiI	Mature heavy chain codon 1 2 3 4 5 6 7
1	MUS IG h 4	TC GCG GCC CAA CCG GCC ATG GCC	GAG GTC CAG CTK CAG CAG TCW G
2	MUS IG h 5	TC GCG GCC CAA CCG GCC ATG GCC	GAG GTG AWG STG GTG GAR TCT G
3	MUS IG h 6	TC GCG GCC CAA CCG GCC ATG GCC	CAG GTY CAG CTG MAG CAG TCT G
4	MUS IG h 7	TC GCG GCC CAA CCG GCC ATG GCC	GAG GTY CAG CTS CAG CAG TCT G
5	MUS IG h 8	TC GCG GCC CAA CCG GCC ATG GCC	GAG GTG AAG CTK RTS GAG TCT G
6	MUS IG h 9	TC GCG GCC CAA CCG GCC ATG GCC	CAG GTG CAG CTK AAG SAG TCA G

primer number	Heavy Chain 3' primers	
7	IgG1 RIL	CTGGTTCGGCCCAACTAGTACAATCCCTGGGCACAATTTTC
8	IgG2a,b RIL	CTGGTTCGGCCCAGATATCACTAGTGGGCCCGCTGGGCTCAA
9	IgG3 RIL	CTGGTTCGGCCCAACTAGTAGAACCTGGGGGGGTACTGG
10	IgM RIL	CTGGTTCGGCCCATCTGCACTAGTTGGAATGGGCACATGCAG

Figure 1B

primer number	Kappa chain primers 5' Primer		NotI		Mature kappa chain codon 1 2 3 4 5 6 7 8
11	MuslG L 1	ATAT	GCGGCCGC	AGGTCTCCTCCTCTTAGCAGCACAACCAGCAATGGCCGAC	ATT STG ATG ACD CAG TCT CCA
12	MuslG L 4	ATAT	GCGGCCGC	AGGTCTCCTCCTCTTAGCAGCACAACCAGCAATGGCCGAT	ATC CAG ATG ACA CAG ACT HCA
13	MuslG L 5	ATAT	GCGGCCGC	AGGTCTCCTCCTCTTAGCAGCACAACCAGCAATGGCCGAT	GTT GTG MTG ACC CAR ACT CCA
14	MuslG L 6	ATAT	GCGGCCGC	AGGTCTCCTCCTCTTAGCAGCACAACCAGCAATGGCCGAC	ATT GTG MTG ACM CAG WCT CCA
15	MuslG L 7	ATAT	GCGGCCGC	AGGTCTCCTCCTCTTAGCAGCACAACCAGCAATGGCCCAA	ATT GTT CTC ACC CAG TCT CCA

Light chain 3' primers

16	CLKAPPA 3'	CATGGCCGGTTGGGCCGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAGGTA-TTTCATTATGACTGTCTCCTTGCTATTAACACTCATTCCTGTTGAAGCTC

Pool amplification primers

17	Light Chain 5'	CTCGCTCGCCCATATGCGGCCGCAGGTCTCCTC
18	Light Chain 3'	CTGGTTCGGCCCACATGGCCGGTTGGGCCGCGA
19	Heavy Chain 5'	CTCGCTCGCCCATCGCGGCCCAACCGGCCATGG
20	Heavy Chain 3'	CTGGTTCGGCCCAAGGCTTACTAGTACAATCCC
21	RIL Heavy Chain 5'	GCGGCCCAACCGGCCATGGCCG
22	RIL Kappa 3'	GGGCCGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAGGTATTTCATTATGACTGTCTCCTTG

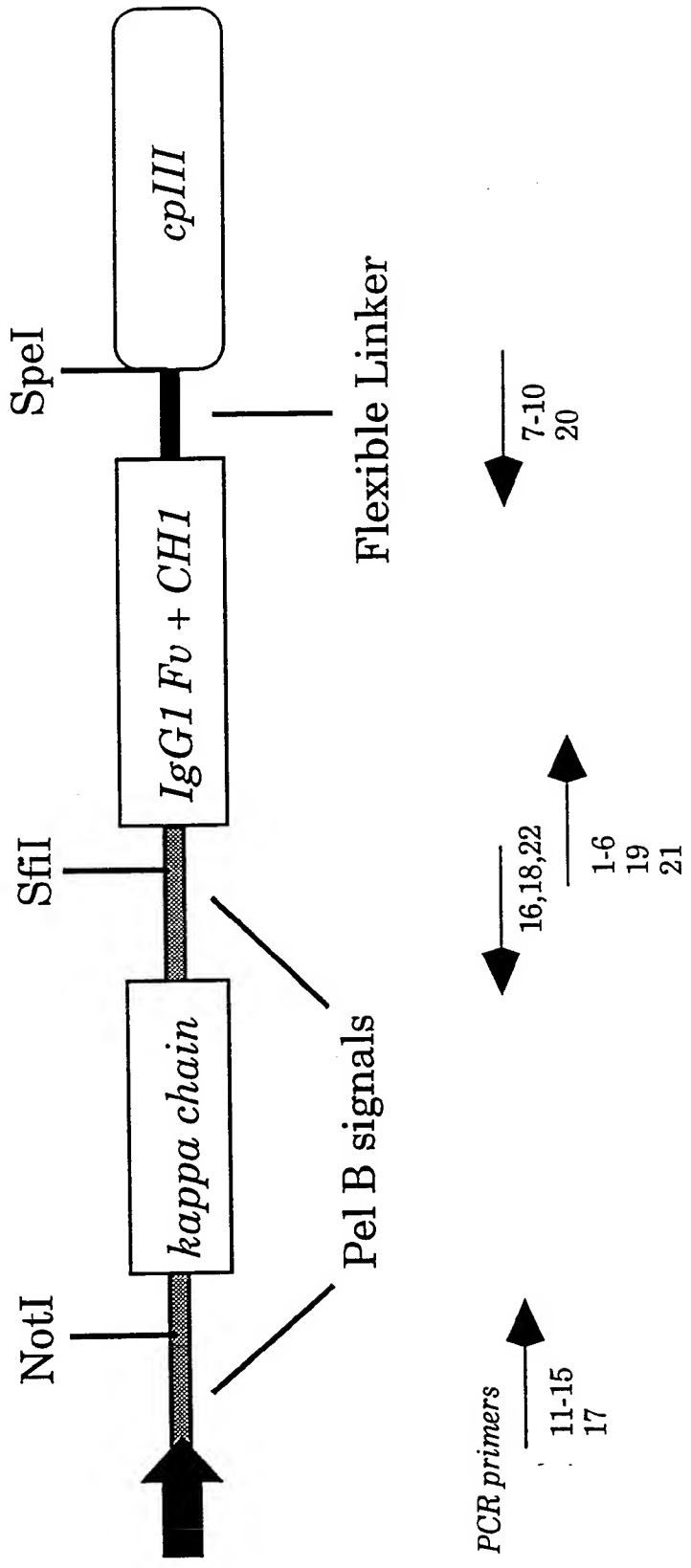
cpIII
SpeI
Flexible Linker
IgG1 Fv + CH1
SfiI
kappa chain
NotI
Pel B signals
7-10
20
16,18,22
1-6
19
21
PCR primers
11-15
17

Figure 2

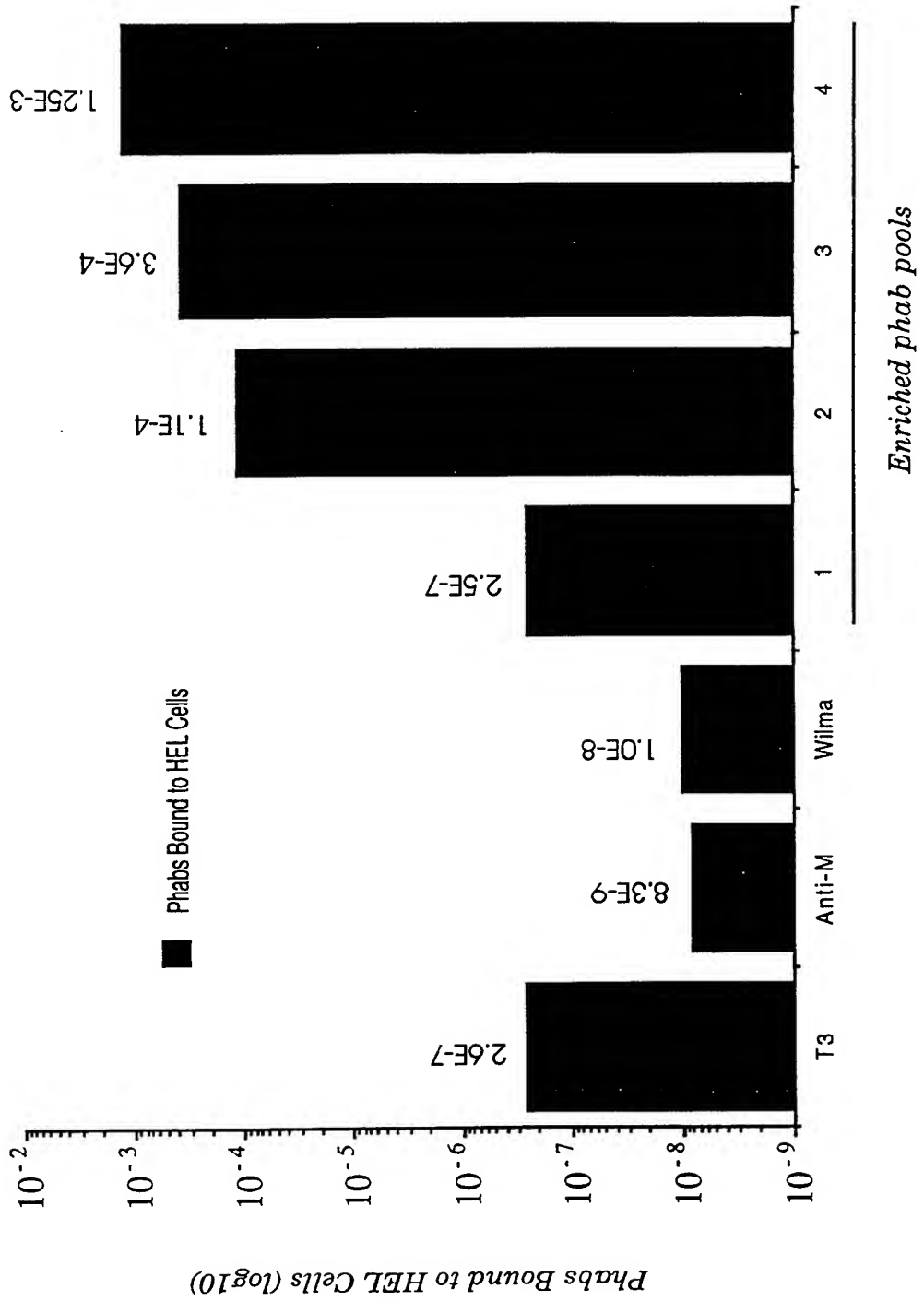
Phabs Bound to HEL Cells (log10)
10-2
10-3
10-4
10-5
10-6
10-7
10-8
10-9
Phabs Bound to HEL Cells
2.6E-7
8.3E-9
1.0E-8
2.5E-7
1.1E-4
3.6E-4
1.25E-3
T3
Anti-M
Wilma
1
2
3
4
Enriched phab pools

Figure 3

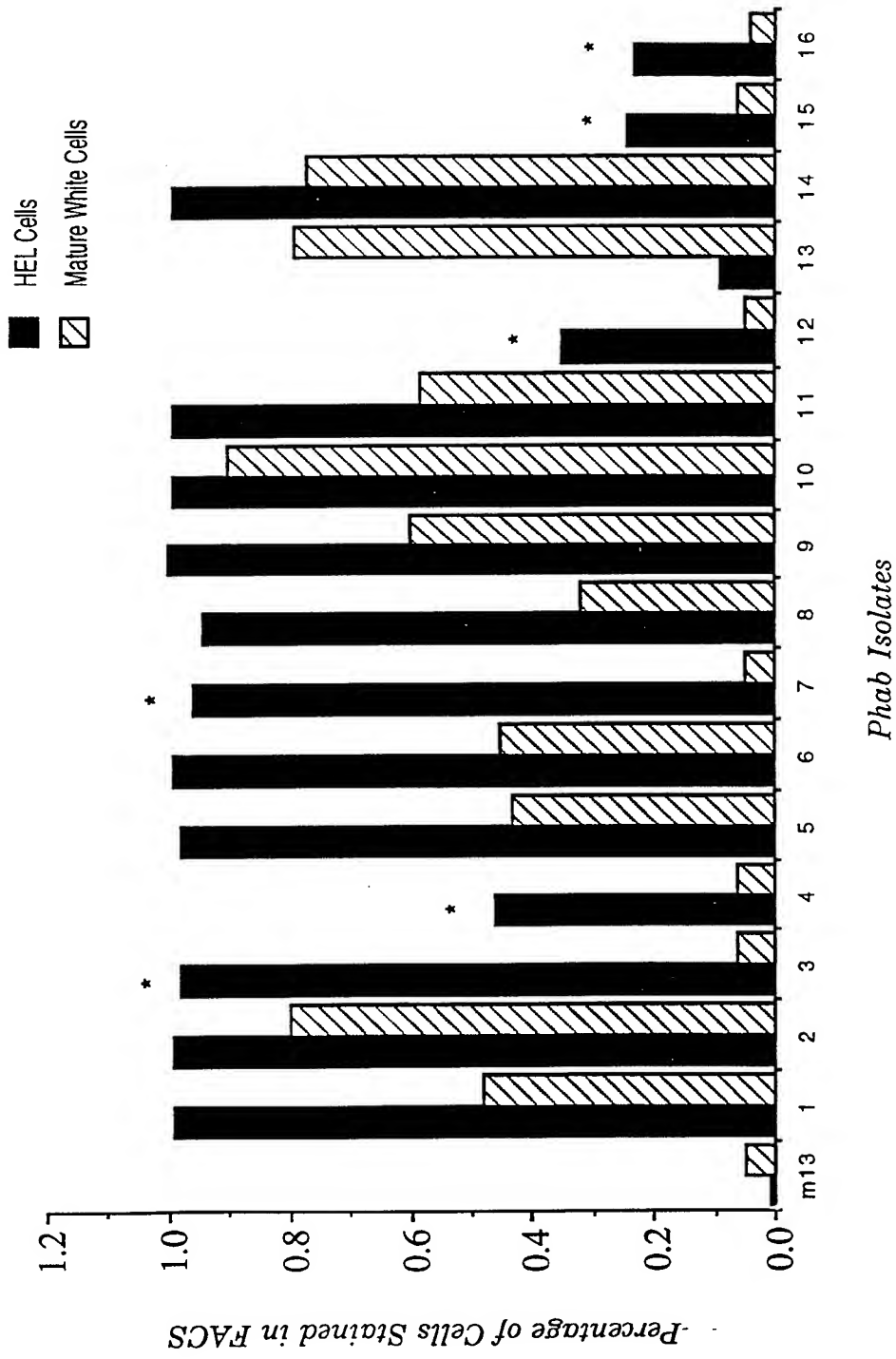
HEL Cells
Mature White Cells
1.2
1.0
0.8
0.6
0.4
0.2
0.0
m13
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
*
Percentage of Cells Stained in FACS
Phab Isolates

Figure 4

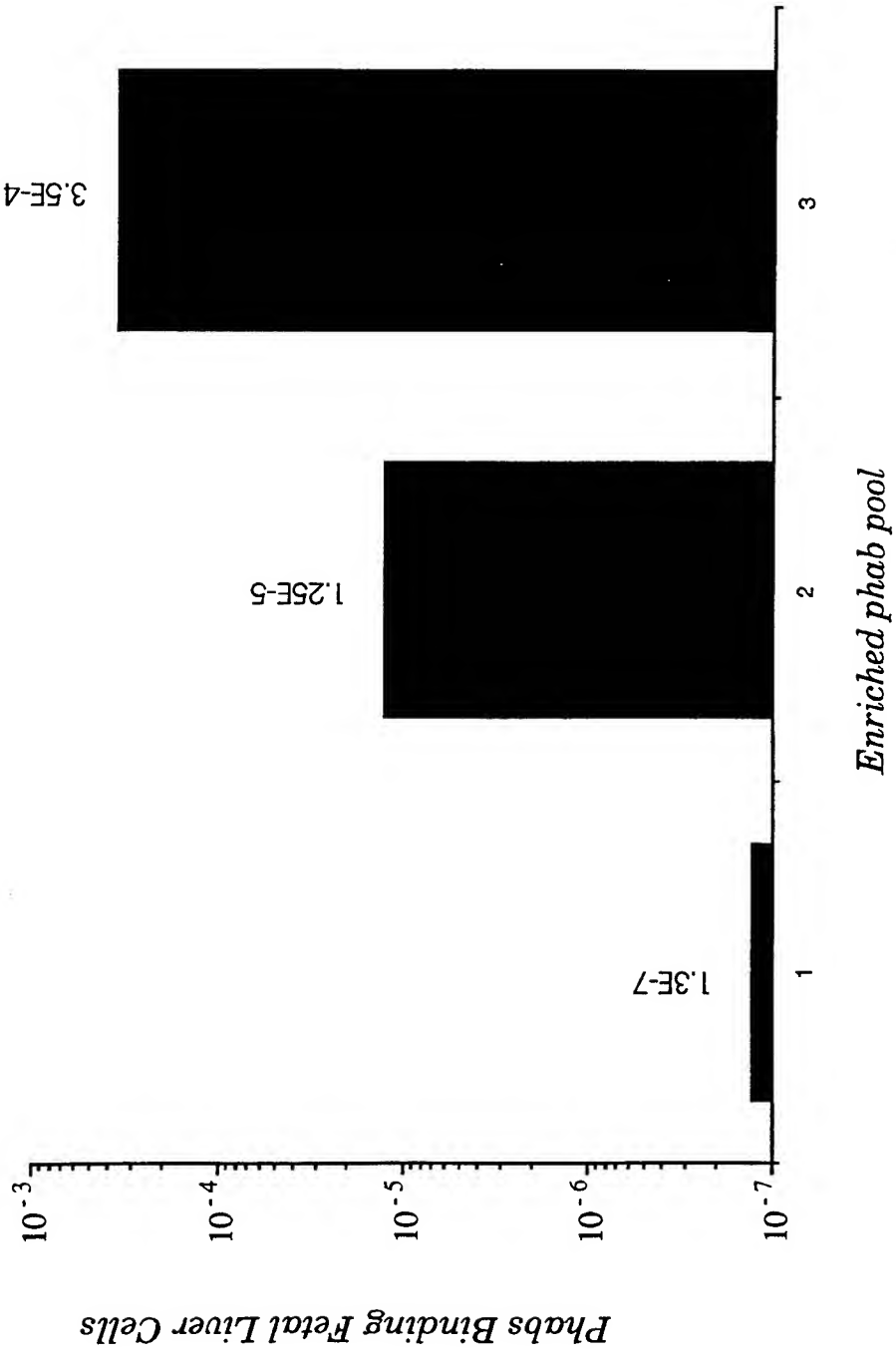
Phabs Binding Fetal Liver Cells
10-3
10-4
10-5
10-6
10-7
1.3E-7
1.25E-5
3.5E-4
1
2
3
Enriched phab pool

Figure 5A

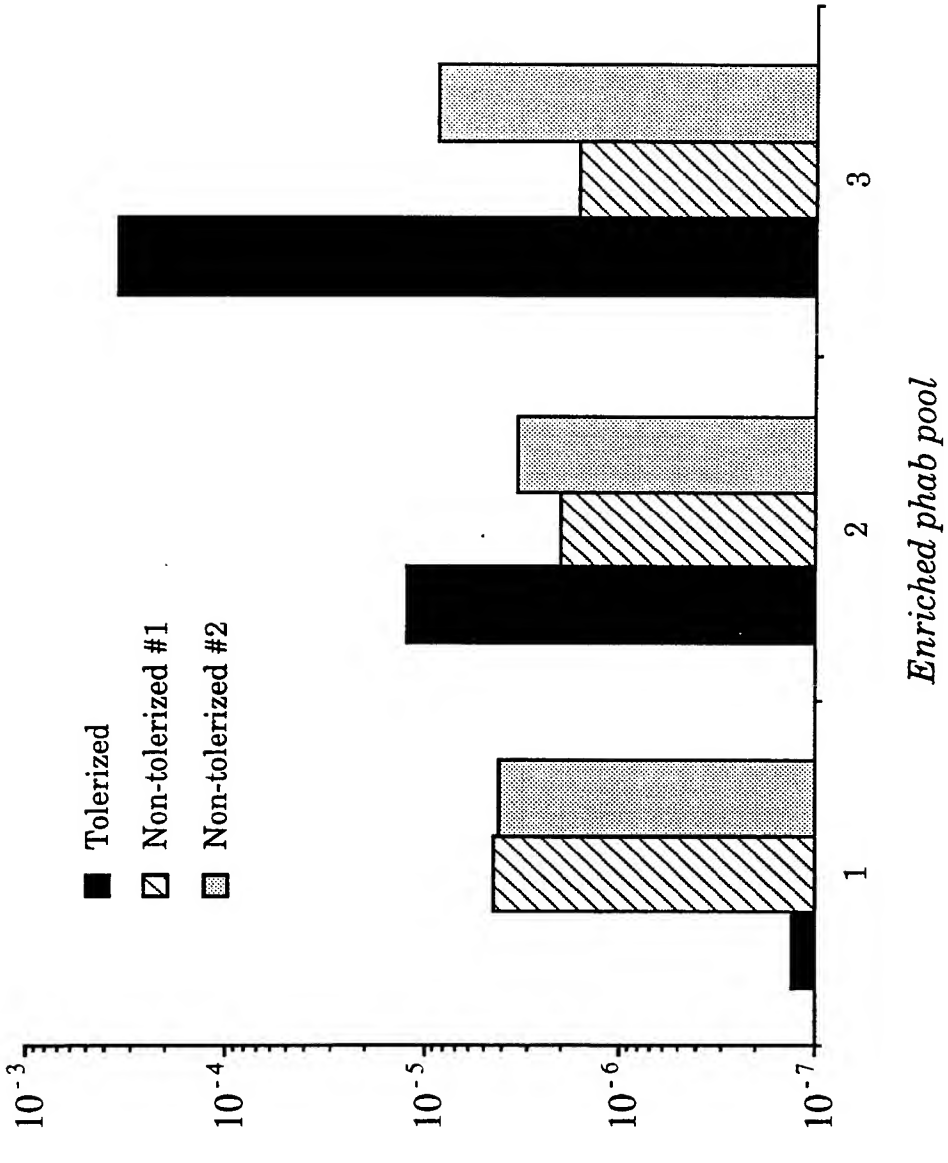
Tolerized
Non-tolerized #1
Non-tolerized #2
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
Phabs Binding Fetal Liver Cells
1
2
3
Enriched phab pool

Figure 5B

Figure 6

Human Immunoglobulin Variable Region PCR Primers

5' End Sense

Human heavy chains

Group A

5'-GGGAATTCATGGACTGGACCTGGAGG(A/G)TC(C/T)TCT(G/T)C-3'

Group B

5'-GGGAATTCATGGAG(C/T)TTGGGCTGA(C/G)CTGG(C/G)TTTT-3'

Group C

5'-GGGAATTCATG(A/G)A(A/C)(A/C)(A/T)ACT(G/T)TG(G/T)-
(A/T)(C/G)C(A/T)(C/T)(C/G)CT(C/T)CTG-3'

Human κ light chain

5'-GGGAATTCATGGACATG(A/G)(A/G)(A/G)(A/G/T)(C/T)CC-
(A/C/T)(A/C/G)G(C/T)GTCA(C/G)CTT-3'

Human λ light chain

5'-GGGAATTCATG(A/G)CCTG(C/G)(A/T)C(C/T)CCTCTC(C/T)-
T(C/T)CT(C/G)(A/T)(C/T)C-3'

3' End sense constant region

Human IgM heavy chain

5'-CCAAGCTTAGACGAGGGGGAAAAGGGTT-3'

Human IgG1 heavy chain

5'-CCAAGCTTGGAGGAGGGTGCCAGGGGG-3'

Human λ light chain

5'-CCAAGCTTGAAGCTCCTCAGAGGAGGG-3'

Human κ light chain

5'-CCAAGCTTTCATCAGATGGCGGGAAGAT-3'

Bases in parentheses represent substitutions at a given residue.

FIGURE 7

CDR3 regions of fetal liver specific phage antibody isolates.

		IgG1				Kappa Chain			
Phab[a]	Isolate	Type	FR3	CDR3[b]	FR4	Type	FR3	CDR3	FR4
1	H3-3	1a	CAR	DPLYGS	WGQ	1a[c]	YFC	SQSTHVLT	FGG
2	Fd2-4					2	FIS	ALKVHM	FSR
3	F4-1	1b	CAR	DPLYGN	WGK	1b	YFC	SQSTHVLT	FGA
4	F3-4					3a	YYC	QQWSSNPPT	FGG
5	Fd4-1, Fd4-16					1c	YFC	SQSHHVLT	FGG
6	F4-11	1c	CIR	DPLYGD	WGQ	1a			
7	F4-7	Va	CAR	GDYGNYGDYFDY	WGQ	4a	YYC	QHSWEIPYT	FGG
8	F3-10,F4-8					4b	YYC	QDSWEIPYT	FGG
9	F4-5					5a	YYC	QQSNEDPYT	FGG
10	Hd3-1					5b	YYC	QQSNEDPFT	FGS
11	F4-16	Vb	CAR	GDYGKYGDYFDH	WGQ	3b	YYC	QQWSSNPPT	FGA
12	Fd3-8	Vc	CAR	GVYGKYGDYFDH	WGK	5a			
13	F3-9					4c	YYC	QHSWEIPFT	FGS
14	Fd3-1,Hd4-1, Fd3-12, Fd3-16	V1	CGGR	EGYGPTGYYSAMDY	WGQ	7	YHC	QQGYSYLT	FGS

[a] F and H indicate enriched on Fetal or HEL cells respectively. The first number indicates the round of enrichment the isolate came from and the second number is the isolate from the indicated round.

[b] This is the hyper variable sequence important for antigen binding and most definitive in defining antibody lineage.

[c] 31Type 1a HC + type 1a kappa chains: F3-1, F3-2, F3-4, F3-6, F4-4, F4-9, F4-12, Fd3-9, Fd3-17, Fd3-20, Hd3-2, Hd3-3, and Fd4 1-16.

Figure 8A

FB3-2 Heavy Chain

```
|-------------FR1--------------| CDR1 |-----FR2-----|      CDR2
EVQLQQSGPE LMMPGASVKI SCKATGYTLS SYWLEWVKQS PGHGLEWIGE ILFGSGSAHY
        A                          S

      |---------------FR3---------------|     CDR3      |-----FR4-
NEKFKGKATF TVDTSSNTAY MQLSSLTSED SAVYYCARGD YGNYGDYFDY WGQGTTLTVS
                                                      QA    F

|   IgG1 CH1 -->
SAKTTPPSVY PLAPGSAAQT NSMVTLGCLV KGYFPEPVTV TWNSGSLSSG VHTFPAVLQS
      S       A                        L

DLYTLSSSVT VPSSTWPSET VTCNVAHPAS STKVDKKIVP RDC
       R
```

FB3-2 Light Chain

```
|----------FR1----------|       CDR1       |------FR2------| CDR2
DIVMTQSPAS LAVSLGQRAT ISCRASQSVS TSRYSYMHWY QQKPGQPAKL LIKFASNLES
                           VR       SH                P       Y

|---------------FR3----------------| CDR3   |---FR4---|  Kappa
GVPARFSGSG SGTDFTLNIH PVEEEDTATY YCQHSWEIPY TFGGGTKLEL KRADAAPTVS
                                                 I

  Constant -->
IFPPSSEQLT SGGASVVCFL NNFYPKDINV KWKIDGSERQ NGVLNSWTDQ DSKDSTYSRS

STLTLTKDEY ERHNSYTCEA THKTSTSPIV KSFNRNEC
```

Figure 8B

H3-3 Heavy Chain

|----------FR1----------| CDR1 |-----FR2-----| CDR2
EVKLMESGGD LVKPGGSLKL SCAASGFTFS DYYMYWVRQT PEKRLEWVAT ISDDGTYTYY

|----------FR3----------| CDR3 |-----FR4--|
ADSVKGRFTI SRDNAKNNLY LQMNSLKSED TAMYYCARDP LYGSWGQGTT LTVSS*AKTTP*

IgG1 CH1 -->
PSVYPLAPGS AAQTNSMVTL GCLVKGYFPE PVTVTWNSGS LSSGVHTFPA VLQSDLYTLS

SSVTVPSSTW SSETVTCNVA HPASSTKVDK KIVPRDC

FB3-2 Light Chain

|----------FR1----------| CDR1 |------FR2------| CDR2
DVVLTQTPLS LPVSLGGQAS ISCRSSQSLV HSNGNTYLHW YLQKPGQSPK LLIYKVSNRF

|----------FR3----------| CDR3 |---FR4---| Kappa
SGVPDRFSGS GSGTDFTLKI SRVEAEDLGV YFCSQSTHVL TFGAGTKLEL K*RADAAPTVS*

Constant -->
IFPPSSEQLT SGGASVVGFL NNFYPKDINV KWKIDGSERQ NGVLNSWTDQ DSKDSTYSRS

STLTLTKDEY ERHNSYTCEA THKTSTSPIV KSFNRNEC